

The Genetic Architecture of Secretory PLA2 (sPLA2) Genes and their Impact on sPLA2 Activity/Mass and Association with CHD Risk

Holly Jane Exeter

University College London

A thesis submitted in accordance with the regulations of the University
College London for the degree of Doctor of Philosophy

Centre for Cardiovascular Genetics

UCL Institute of Cardiovascular Science

For Mum and Dad

Declaration

I, Holly Jane Exeter confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. In relation to specific chapters I would like to clarify my role:

In **Chapter 3** I was responsible for identifying SNPs of interest from the results of previously genotyped studies. I was then responsible for the in depth literature and bioinformatics study undertaken for *PLA2G2A*, sPLA2-IIA and the SNPs of interest. I designed the primers for all cloning and EMSA work and carried out all of the cloning experiments, luciferase assays, TaqMan Gene Expression assays and EMSAs according to the methods described. I also carried out the statistical work relating to the luciferase assay experiments.

In **Chapter 4** I was responsible for genotyping 2 of the studies in the final meta-analysis; IMPROVE and CYPRUS. I used the PRISMA guidelines to compile an in depth literature search of the sPLA2 inhibitor, varespladib. I was part of the core group of investigators who met to discuss the results and direction of the study (participants of the MR study are at the end of the thesis).

In **Chapter 5** I was responsible for identifying SNPs of interest from the results of previously genotyped studies. I was then responsible for the in depth literature and bioinformatics study undertaken for *PLA2G5*, sPLA2-V and the SNPs of interest. I genotyped SNPs as indicated in Table 5.1.

In **Chapter 6** I was responsible for identifying the SNP of interest from the results of previously genotyped and published studies. I was then responsible for the in depth literature and bioinformatics study undertaken for *PLA2G10*, sPLA2-X and the SNP of interest. With Jutta Palmen I designed the RFLP protocol for genotyping rs4003228 and rs4003232. With Dr Montse Guardiola I was responsible for genotyping rs4003232 and rs4003228 in all studies.

Regarding the Advanced Study of Aortic Pathology (ASAP) (the Karolinska Institute, Sweden). All analysis was carried out by Dr. Lasse Folkersen, based on ideas conceived by myself, Prof. Philippa Talmud, Dr. Michael Holmes and Dr. Lasse Folkersen.

Holly J. Exeter

Acknowledgements

There are many people who deserve my thanks and acknowledgement for their roles in this thesis. First I would like to express my sincere and heartfelt thanks to my supervisor, Philippa Talmud. Your honesty and openness throughout this process have been vital and your faith in me and excellent guidance has driven me to bring out the best in myself. Thank you. I would also like to convey my gratitude to Professor Steve Humphries, whose generously given time and expertise have been invaluable, and to Professor Aroon Hingorani for his encouragement and advice throughout.

The research carried out in this thesis was funded by the British Heart Foundation and I am extremely grateful to them for the financial support that they have provided as well as their keen interest in the progress of this work.

CVG is not an everyday lab. It has been a pleasure to be part of such a diverse and exciting group, but more than that I have had enormous fun and made some fantastic friends. The camaraderie between all staff is fantastic and the depth of expertise on every kind of subject has meant that I have never been short of conversation or someone willing to help! Special thanks have to go to those that have shared the PhD office with me; Melissa, Donald, Klelia, Angus and Marta, you made it a pleasure to come to work and have provided tea and sympathy and shared laughs and excitement in equal measure! To those in the RA office, Jacquie, Jay, Wendy, Ros, Phil, Dauda and KaWah, you have all been instrumental in helping me to problem solve and learn new skills as well as providing a warm and friendly environment to come and work with. Thanks to all of you.

Of all the group at CVG I need to single out Jutta Palmen who has been a great friend and an outstanding mentor in the lab. Her patience has allowed me to steadily build my confidence in my abilities and vastly improve my lab skills. I have experienced many ups and downs during the course of my thesis, but I have consistently been able to trust Jutta and Philippa especially to support and guide me to find the right solutions.

Thanks are extended also to those who I have collaborated with closely throughout my studies; Dr Michael Holmes, Dr Lasse Folkersen, Dr Jackie Cooper and Dr Montse Guardiola. Jackie and Michael for their statistical expertise, Lasse for access to and analysis of his groups' study, the Advanced Study of Aortic Pathology (ASAP) and Montse for her collaboration on the *PLA2G10* genotyping data. These individuals and their contributions are indicated throughout the thesis.

Outside of science I have the most amazing group of friends. The support and tolerance they have provided me with over many years, but especially during my PhD studies, has been fantastic; Hannah, Emily, EJ, Averil, Carla, Jasmine, Danielle and Mellissa, you're all brilliant! My special thanks and gratitude go to my oldest friend Lauren, who has lifted and motivated me on more occasions than I can say. Thank you to all of you for listening (and for being amazing dance partners!).

Thanks to my dear family, all of you, my aunts, uncles, cousins, Grandma, Nan, Granddad and Granddad, you have all had an indefinable influence on my life and I greatly appreciate all that you have done for me.

Alice and Charlie, my superb siblings(!) You put up with me day in and day out and have lived through all of my many stressful moments! Your tolerance is hugely appreciated and your capacity to make me laugh (intentionally or not!) and cheer me up has been greatly needed! We are an awesome trio!

Finally, I would not be anything without the limitless love and support of my wonderful parents. Mum and Dad, your help and encouragement have shaped me and led me to believe that I could be whatever I wanted to be. There is no way for me to fully express my gratitude for your love and commitment- both emotional and financial over all these years, but I hope I've made you proud. Thank you both for being you!

Abstract

Secretory phospholipase A2 group (sPLA2) enzymes hydrolyze the sn-2 ester bond of phospholipids, producing lysophospholipids and free fatty acids. sPLA2s have been identified as biomarkers of atherosclerosis in observational and animal studies.

The aim of this study was to identify functional variants in *PLA2G2A* and, to a lesser extent, *PLA2G5* and *PLA2G10* (sPLA2 encoding genes). These variants could then serve as tools to examine their contribution to sPLA2-activity, since this is a composite measure of sPLA2-IIA,-V and -X, and associations with CHD risk/CHD traits.

Two *PLA2G2A* SNPs were identified as being associated with sPLA2-IIA mass/sPLA2 activity. Luciferase assays, EMSA and RNA expression were used to examine their allelic differences. Rs3767221G showed ~55% lower luciferase activity compared to rs3767221T ($p=1.22\times 10^{-35}$), and stronger EMSA binding of a nuclear protein compared to the T-allele. For rs11573156 C >G there were no luciferase or EMSA allelic differences. For rs11573156 there was no differential allelic lymphocyte cDNA expression for exons 5-6, but G-allele carriers ($n=7$) showed a trend to lower exon 1-2 expression versus CC individuals. In the ASAP study ($n=223$), a SNP acting as a proxy for rs11573156 ($r^2=0.91$) on the expression array showed allelic difference of ~25% in liver expression of total *PLA2G2A* (1.67×10^{-17}). However, exon 2 specific expression was greatly reduced (4.5×10^{-5}) compared to exons 3-6 (10^{-10} to 10^{-20}), implying allele-specific exon 2 skipping as the functional basis of rs11573156 association. Rs11573156 was used as a genetic tool in a large Mendelian randomisation (MR) collaboration to establish the causality of sPLA2-IIA for CHD. While the sPLA2-IIA mass was strongly associated with CHD events, rs11573156 was not associated with outcome. Similar approaches were taken for studying *PLA2G5* and *PLA2G10*, using eQTL data and smaller meta-analyses.

Both *PLA2G2A* SNPs are functional. Using rs11573156 as a genetic instrument, MR suggested that sPLA2-IIA is not causally associated with CHD. Based on initial meta-analyses, neither *PLA2G5* nor *PLA2G10* appeared to be associated significantly with sPLA2 activity in cardiovascular related tissues. This conclusion is consistent with the outcome of the recent Phase III trial for the sPLA2 inhibitor varespladib, which was terminated due to lack of efficacy since the drug was shown to be having no significant effect on the reduction of secondary cardiovascular events.

Contributors to the Mendelian Randomisation Study of sPLA2-IIA and CHD

Michael V Holmes, MRCP, MSc

Tabassome Simon, MD, PhD

Lasse Folkersen, PhD,

Folkert W Asselbergs, MD, PhD

Montse Guardiola, PhD

Jackie A Cooper, MSc

Jutta Palmen,

Jaroslav A Hubacek, DSc, PhD

Kathryn F Carruthers, MPhil

Benjamin D Horne, PhD, MPH

Kimberly D. Brunisholz, MST

Jessica L Mega, MD, PhD

Erik PA van Iperen, MSc

Mingyao Li, PhD

Maarten Leusink, MSc

Stella Trompet, PhD,

Jeffrey JW Verschuren, MD, MSc

G. Kees Hovingh, MD, PhD

Abbas Dehghan, MD, PhD

Christopher P Nelson, PhD

Salma Kotti, PharmD, PhD

Nicolas Danchin, MD

Markus Scholz, PhD

Christiane L Haase, PhD

Dietrich Rothenbacher, MD

Daniel I Swerdlow, PhD

Karoline B Kuchenbaecker, Dipl.

Psych.,MSc

Ferdinand van 't Hooft, MD, PhD

Karl Gertow, PhD

Ulf de Faire, MD, PhD

Andrie G Panayiotou, PhD

Elena Tremoli, PhD

Damiano Baldassarre, PhD

Daniel Teupser, MD, PhD

Fabrizio Veglia, PhD

Lesca M. Holdt, MD, PhD

Frank Beutner, MD, PhD

Ron T Gansevoort, MD, PhD

Gerjan J Navis, MD, PhD

Irene Mateo Leach, PhD

Lutz P. Breitling, MD

Hermann Brenner, MD, MPH

Dhayana Dallmeier, MD

Anders Franco-Cereceda, MD, PhD

Jolanda MA Boer, PhD

Jeffrey W Stephens, PhD, FRCP

Marten H Hofker, PhD

Alain Tedgui, PhD

Albert Hofman, MD, PhD

André G Uitterlinden, PhD

Vera Adamkova, MD, PhD

Jan Pitha, MD, PhD

N. Charlotte Onland-Moret, PhD

Maarten J Cramer, MD, PhD

Hendrik M. Nathoe, MD, PhD

Wilko Spiering, MD, PhD

Olaf H Klungel, PharmD, PhD

Meena Kumari, PhD

Peter H Whincup, FRCP

David A Morrow, MD

Peter S Braund, MSc

Alistair S Hall, FRCP, PhD

Pieter A. Doevendans, MD, PhD

Mieke D. Trip, MD, PhD

Martin D Tobin, PhD, FFPH

Anders Hamsten, FRCP, PhD

Wolfgang Koenig, MD

Andrew N Nicolaides, FRCS, PhD

Ian N M Day, FRCPath
John F Carlquist, PhD
Tom R Gaunt, PhD
Ian Ford, PhD
Naveed Sattar, FRCP
Debbie A Lawlor, PhD
Richard W Morris, PhD
Manjinder S. Sandhu, PhD
Rudolf Poledne, PhD
Anke H. Maitland-van der Zee, PhD
Kay-Tee Khaw, FRCP
Brendan J Keating, PhD
Pim van der Harst, MD, PhD
Jackie F Price, FFPH, MD
Jaqueline C M Witterman, PhD
Oscar H Franco, MD, PhD
J Wouter Jukema, MD, PhD
Peter de Knijff, PhD
Anne Tybjaerg-Hansen, MD DMSc
Daniel J Rader, MD
Nilesh J Samani, MD
Mika Kivimaki, PhD,
Keith A A Fox, FRCP
Steve E Humphries, FRCPath, FRCP, PhD
Jeffrey L Anderson, MD
S Matthijs Boekholdt, MD, PhD
Per Eriksson, PhD
Aroon D Hingorani, FRCP, PhD
Marc S Sabatine, MD, PhD
Ziad Mallat, MD, PhD
Juan P Casas, MD, PhD
Philippa J Talmud, FRCPath, DSc

Table of Contents

Declaration	3
Acknowledgements	4
Abstract	6
Contributors to the Mendelian Randomisation Study of sPLA2-IIA and CHD	7
Chapter Contents	9
List of Figures	17
List of Tables	22
List of Acronyms	24
Publications Resulting From This Thesis	28

Contents

Chapter 1 Introduction

1.	Cardiovascular disease: A Worldwide Problem	29
1.1.	Coronary Heart Disease and Atherosclerosis	29
1.2.	Pathology of Atherosclerosis	31
1.2.1.	The Morphology of a Normal Artery	31
1.2.2.	Development of Atherosclerosis	32
	The Step-Wise Progression of an Atherosclerotic Plaque	33
	Plaque Initiation by Endothelial Dysfunction	33
	Development of the Plaque: LDL Accumulation and Modification	34
	Foam Cell Formation	34
	Plaque progression: Necrotic Core Formation and Plaque Rupture	35
1.3.	Predicting the Risk of CHD	37
1.3.1.	Some known and Emerging Risk Factors of CHD	37

1.3.2.	Environmental Risk Factors	39
1.3.3.	Factors with a Strong Genetic Component	39
	Inflammatory Markers of Atherosclerosis	40
	Diabetes and Glucose Control as a CHD Risk Factor	41
1.4.	Lipoprotein Metabolism	42
1.4.1.	Structure of Lipoproteins	42
1.4.2.	Phospholipids	44
1.4.3.	Metabolism of Lipoproteins	44
	Transport of Dietary Fat	45
	Transport of Endogenous Fat	45
	Reverse Cholesterol Transport and the Role of HDL	45
1.4.4.	Lipoprotein Metabolism as a Risk Factor	46
1.5.	The Secretory Phospholipase A2 Family	47
1.5.1.	Secretory phospholipase A2 Group IIA (sPLA2-IIA)	54
1.5.2.	Secretory phospholipase A2 Group V (sPLA2-V)	55
1.5.3.	Secretory phospholipase A2 Group X (sPLA2-X)	56
1.5.4.	The Genetics of Secretory Phospholipases A2s	60
1.5.5.	Measuring sPLA2 Mass and sPLA2 Activity	61
1.5.6.	Biological Functions of the sPLA2 Family	62
1.5.7.	sPLA2s in Disease	63
	sPLA2-IIA Role in Rheumatoid Arthritis	64
	sPLA2-IIA Role in Atherosclerosis and CHD	65
	Observational Data	65
	Animal Models	68
	Suggested sPLA2-IIA Pathway to Atherosclerosis	68
	sPLA2-V in Arthritis and Respiratory Disease	69
	sPLA2-V in Atherosclerosis and CHD	70
	sPLA2-X in Respiratory Disease	70

	sPLA2-X in Atherosclerosis and CHD	72
1.5.8.	An Overview of sPLA2 Actions	72
	Summary of the sPLA2 Association with CHD Based on Current Data	73
1.6.	Varespladib: An sPLA2 Inhibitor Under Trial	76
1.6.1.	Varespladib Structure and Production	77
1.7.	Genetic Variations and Disease	81
1.7.1.	Genome-Wide Association Scan	82
1.7.2.	Indirect/Direct Candidate Gene Approach	82
1.7.3.	Mendelian Randomisation	86
1.8.	Hypothesis	89
1.9.	Aims	90

Chapter 2 Materials and Methods

2.1.	Commonly Used Reagents, Stocks, Products and Kits (alphabetical)	91
2.2.	Commonly Used Websites, Software and Services (alphabetical)	96
2.3.	Commonly Used Laboratory Machines and Equipment (alphabetical)	98
2.4.	Standard Methods	99
2.4.1.	SNP Selection and Bioinformatics	99
	Analysing the sPLA2 Inhibitor, Varespladib	100
2.4.2.	DNA Standardisation and Determining DNA Concentration	100
2.4.3.	Primer Design	100
2.4.4.	Polymerase Chain Reaction (PCR)	100
2.4.5.	Agarose Gel Electrophoresis	101
2.4.6.	Genotyping	102
2.4.6.1.	TaqMan Genotyping	102
	TaqMan Genotyping Signal Detection	105
	TaqMan Genotyping <i>PLA2G2A</i> SNPs of Interest	106
	TaqMan Genotyping <i>PLA2G5</i> SNPs of Interest	107

	Genotyping <i>PLA2G10</i> SNPs of Interest	107
	Micro Array Diagonal Gel Electrophoresis	109
2.4.7.	Luciferase Assay Analysis	111
2.4.7.1.	Isolating a DNA Fragment for Cloning into a pGL3-Luciferase Reporter Vector	111
2.4.7.2.	Ligation of Digested Vector and DNA Fragment (Insert)	111
2.4.7.3.	Transforming Vector/Insert Complex into Competent Cells	112
2.4.7.4.	Amplification of Transformed Colonies and DNA Purification by Miniprep	112
	Plasmid Design for rs11573156 C>G	112
	Plasmid Design for rs3767221 A>C	113
2.4.7.5.	Verification of Amplification of the Correct DNA Sequence	114
	<i>PLA2G2A</i> rs11573156 Site Directed Mutagenesis to Introduce a Base	117
2.4.7.6.	Amplification of Re-grown Colonies and DNA Purification by Maxiprep	117
2.4.7.7.	Transfection of Purified Plasmid into a Human Cell Line	118
2.4.7.8.	Cell Lysis and β -Gal Assay	119
2.4.7.9.	Luciferase Assay	119
2.4.8.	LightShift Chemiluminescent Electrophoretic Mobility Shift Assay (EMSA)	120
2.4.8.1.	Designing, Biotin Labelling and Annealing Primers	120
	EMSA Probes for the Analysis of rs11573156	120
	EMSA Probes for the Analysis of rs3767221	121
2.4.8.2.	Nuclear Extraction	121
2.4.8.3.	Polyacrylamide Gel (6%) Preparation	122
2.4.8.5.	EMSA Blotting Membrane	123
2.4.8.6.	LightShift Chemiluminescent EMSA Detection System	123
2.4.8.7.	EMSA Visualisation	124
2.4.9.	TaqMan Gene Expression	124
2.4.9.1.	RNA Extraction	124
2.4.9.2.	Complementary DNA (cDNA) Synthesis by Reverse Transcription	124

2.4.9.3.	TaqMan Gene Expression Assays (20x)	125
	TaqMan Gene Expression Method Description and Conditions	125
	TaqMan Gene Expression Signal Detection	128
	TaqMan Gene Expression Analysis for rs11573156 C>G	131
2.4.10.	The Advanced Study of Aortic Pathology (ASAP) Gene Expression Analysis	132
2.4.11.	Statistics	133
2.5.	Meta Analyses Study Populations and Specifications	134
2.5.1.	Mendelian Randomisation Using <i>PLA2G2A</i> rs11573156	134
	Mendelian Randomisation Method	138
	<i>PLA2G5</i> Study Populations Used to Analyse SNPs of Interest	138
	<i>PLA2G10</i> Study Populations Used to Analyse SNPs of Interest	140
	The DRD Cohort	141

Chapter 3 Identification of Functional Variants of *PLA2G2A*

3.	Chapter Background	142
3.1.	Chapter Introduction	142
3.2.	Results	144
3.2.1.	Bioinformatics Study of <i>PLA2G2A</i>	144
3.2.2.	Bioinformatics Study of rs11573156	147
3.2.3.	Bioinformatics Study of rs3767221	158
3.2.4.	Functional Experiments	165
3.2.4.1.	Luciferase Assay Results for Analysis of rs11573156	165
3.2.4.2.	Luciferase Assay Results for Analysis of rs3767221	165
3.2.5.	EMSA Results	166
3.2.5.1.	EMSA Results for the Analysis of rs11573156	166
3.2.5.2.	EMSA Results for the Analysis of rs3767221	168
3.2.6.	TaqMan Gene Expression Results	170
3.2.6.1.	TaqMan Gene Expression Results for the Analysis of rs11573156	170

3.2.7.	ASAP Study Gene Expression Analysis	173
	Bioinformatics Study of rs10732279	173
3.2.7.2.	ASAP Study Gene Expression Analysis for rs11573156 Proxy	173
3.2.7.3.	ASAP Study Gene Expression Analysis for rs3767221	177
3.3.	Discussion	178
3.3.1.	Rs11573156 C>G as a Functional SNP	178
3.3.2.	Rs3767221 A>C as a Functional SNP	180
3.4.	Summary of Findings	181
3.4.1.	Summary of the Potential Functionality of rs11573156	181
3.4.2.	Summary of the Potential Functionality of rs3767221	182
3.5.	Conclusion	182

Chapter 4 Using Genetics to Test the Causality of sPLA2-IIA for CHD: A Mendelian Randomisation Approach

4.	Chapter Background	183
4.1.	Chapter Introduction	183
4.1.1.	Formation of the sPLA2-IIA Mendelian Randomisation Study	183
4.1.2.	The Effects of sPLA2-IIA Levels/sPLA2 Activity on Cardiovascular Risk Traits	184
4.1.3.	Varespladib Literature Study	184
4.1.4.	Identifying the Strongest Genetic Instrument	188
4.2.	Results	188
4.2.1.	Rs11573156 C>G the Strongest and Most Specific Genetic Instrument	188
4.2.2.	Results for Observational Analysis of sPLA2-IIA and Myocardial Infarction	198
4.2.3.	Mendelian Randomisation Results: rs11573156 and Cardiovascular Outcomes	204
4.3.	Discussion	207
4.4.	Conclusion	209

Chapter 5

The Potential Contribution of sPLA2-V to sPLA2 activity and Risk of CHD

5.1.	Chapter Introduction	210
5.1.1.	Background: sPLA2 Isoenzyme Analysis	210
5.1.2	Comparing sPLA2-V and sPLA2-IIA	211
5.1.3.	The Potential Contribution of sPLA2-V to sPLA2 Activity	211
5.1.4.	Investigating the Contribution of sPLA2-V to CHD Risk	212
5.1.5.	SNP Selection	213
	Bioinformatics of tSNPs	215
5.2.	Results	215
5.2.1.	Bioinformatics Analysis	215
5.2.1.1.	Bioinformatics Analysis of <i>PLA2G5</i>	215
5.2.1.2.	Bioinformatics Analysis of the Lead tSNP rs11573185 C>A	215
5.2.1.3.	Bioinformatics Analysis of rs525380 A>C	219
5.2.2.	<i>PLA2G5</i> mRNA Expression in the ASAP Study	221
5.2.3.	Association of <i>PLA2G5</i> SNPs with <i>PLA2G2A</i> mRNA Levels	226
5.2.4.	Combined Meta-Analysis Data for rs11573185	227
5.2.5.	Combined Meta-Analysis Data for rs525380	232
5.3.	Discussion	234
5.4.	Conclusion	237

Chapter 6 The Potential Contribution of sPLA2-X to sPLA2 Activity

6.1.	Chapter Introduction	238
6.1.1.	sPLA2-X Background	238
6.1.2.	Identifying Tagging-SNPs of Interest	238
6.2.	Results	239
6.2.1.	Bioinformatics Analysis of <i>PLA2G10</i>	239
6.2.2.	Results from EPIC-Norfolk and GRACE-France	241

6.2.3.	Bioinformatics Analysis of rs4003228 (R38C)	243
6.2.4.	Meta-Analysis for rs4003228 (R38C)	243
6.2.5.	<i>PLA2G10</i> mRNA Expression: The ASAP Study	246
6.3.	Discussion	247

Chapter 7 Discussion

7.1.	General Discussion	251
7.1.1.	Observational Data	251
7.1.2.	Animal (<i>in vivo</i>) Models	252
7.1.3.	<i>In vitro</i> Analysis	253
7.1.4.	Mendelian Randomisation	254
7.1.5.	Inhibitor Drug Design	256
7.1.6.	ASAP Study: The Value of eQTLs	256
7.2.	Overall Summary of Results	257
7.2.1.	Limitations	261
7.3.	Conclusions and Further Work	261

Bibliography	263
---------------------	-----

List of Figures

Figure 1.1.	WHO 10 Leading Causes of Death (2008) Low Income Countries	30
Figure 1.2.	WHO 10 Leading Causes of Death (2008) High Income Countries	30
Figure 1.3.	The Morphology of the Normal Artery	32
Figure 1.4.	The Development of an Atherosclerotic Plaque	36
Figure 1.5.	The Structure of a Circulating Plasma Lipoprotein	43
Figure 1.6.	An Overview of the Lipoprotein Metabolism Process	44
Figure 1.7.	The Site of sPLA2 hydrolysis	48
Figure 1.8.	The Potential Outcomes of sPLA2 Hydrolysis	49
Figure 1.9.	The sPLA2 Isoforms	51
Figure 1.10	The Specific Catalytic Dyad Active Site of sPLA2 Enzymes	52
Figure 1.11.	The Protein Coding Sequences for sPLA2-IIA, V and X	57
Figure 1.12	The Heparan Sulphate Pathway and the External Plasma Membrane Pathway	58
Figure 1.13	The <i>PLA2</i> Gene Cluster Chromosome 1	60
Figure 1.14	<i>PLA2G10</i> Chromosome 16	61
Figure 1.15	A Schematic Representation of sPLA2 Pathways	73
Figure 1.16	The Potential Contribution of sPLA2 Enzymes to Atherosclerosis	75
Figure 1.17	The Structure of Varespladib, an sPLA2 Inhibitor	78
Figure 1.18	A Schematic Representation of <i>PLA2G2A</i>	84
Figure 1.19	A Schematic Representation of <i>PLA2G5</i>	85
Figure 1.20	Mendelian Randomisation versus an RCT	88
Figure 2.1.	The pGL3 Basic Luciferase Reporter Vector	94
Figure 2.2.	The pGL3 Promoter Luciferase Reporter Vector	95
Figure 2.3.	The pGL3-Control Vector	95
Figure 2.4.	A Typical PCR Reaction	104

Figure 2.5.	A Typical Allelic Discrimination Plot	106
Figure 2.6.	A Typical RFLP MADGE Gel Analysis	110
Figure 2.7.	The TaqMan Gene Expression PCR Process	127
Figure 2.8.	The Normalisation of TaqMan Gene Expression using ROX	129
Figure 2.9a.	A Typical Amplification Curve for TaqMan Gene Expression	130
Figure 2.9b.	A Typical TaqMan Gene Expression Results Plot	131
Figure 3.1.	Known <i>PLA2G2A</i> Transcripts	146
Figure 3.2.	An LD Heatplot for <i>PLA2G2A</i> Tagging-SNPs	147
Figure 3.3.	UCSC Genome Browser Data for <i>PLA2G2A</i> rs11573156	150
Figure 3.4.	An Example of Alternate Splicing	152
Figure 3.5.	An LD Manhattan Plot for rs3767221	159
Figure 3.6.	UCSC Genome Browser Data for <i>PLA2G2A</i> rs3767221	162
Figure 3.7.	Luciferase Assay Results for rs11573156	165
Figure 3.8.	Luciferase Assay Results for rs3767221	166
Figure 3.9.	EMSA Image for rs11573156	167
Figure 3.10.	EMSA Image for rs3767221	168
Figure 3.11.	EMSA Image for rs3767221 Comparing Specific Binding	169
Figure 3.12.	TaqMan Genotyping Results for rs11573156 in the DRD Cohort	171
Figure 3.13	TaqMan Gene Expression Assay Results for rs11573156	172
Figure 3.14	A Schematic of <i>PLA2G2A</i> and relevant Tagging-SNPs	173
Figure 3.15	Expression of <i>PLA2G2A</i> mRNA in the ASAP Study	174
Figure 3.16	Association of rs10732279 and <i>PLA2G2A</i> mRNA in the Liver	175
Figure 3.17	Genotype and Exon Specific <i>PLA2G2A</i> mRNA Expression for rs10732279	176
Figure 3.18	Differential <i>PLA2G2A</i> mRNA Expression Predicted for rs3767221 In the Liver	177
Figure 4.1.	Results of the PRISMA Literature Analysis for Varespladib	185
Figure 4.2A.	Hardy-Weinberg Equilibrium Equation for rs11573156 In EPIC-Norfolk	190

Figure 4.2b.	Hardy-Weinberg Equilibrium Equation for rs11573156 In GRACE-France	190
Figure 4.2c.	Hardy-Weinberg Equilibrium Equation for rs3767221 In EPIC-Norfolk	191
Figure 4.2d.	Hardy-Weinberg Equilibrium Equation for rs3767221 In GRACE-France	191
Figure 4.3.	The Effect of rs11573156 Variants on sPLA2-IIA Mass and sPLA2 Activity	196
Figure 4.4.	Comparing mRNA Expression of <i>PLA2</i> Genes Differentially Associated with SNPs from the ASAP Arrays	197
Figure 4.5.	Hardy-Weinberg Equilibrium Equation for rs11573156 In GRACE Scotland	199
Figure 4.6.	Hardy-Weinberg Equilibrium Equation for rs11573156 In UDACS	200
Figure 4.7.	The Frequency Distribution of sPLA2-IIA mass	203
Figure 4.8.	Association Between rs11573156 and Major Vascular Events	205
Figure 4.9.	Mendelian Randomisation Analysis Results for rs11573156 and CHD	206
Figure 5.1.	The <i>PLA2G5</i> Gene and Previously Identified Tagging-SNPs	212
Figure 5.2.	The Chromosomal Location of rs1573185	216
Figure 5.3.	An LD Heatplot Showing rs11573185 D' with Other SNPs in the Region	217
Figure 5.4.	An LD Plot Showing rs11573185 R^2 with Other SNPs in the Region	218
Figure 5.5.	An LD Plot Showing rs525380 R^2 with Other SNPs in the Region	220
Figure 5.6.	UCSC Genome Browser Data for <i>PLA2G5</i> tagging SNPs	221
Figure 5.7.	The Association of <i>PLA2G5</i> SNPs with <i>PLA2G5</i> mRNA in ASAP	222
Figure 5.8.	Differential <i>PLA2G5</i> mRNA Expression Associated with rs11573185 Measured in all ASAP Study Tissues	223
Figure 5.9.	The Per Allele Effect of rs11573185 on <i>PLA2G5</i> mRNA Expression in the Aorta Adventitia	224

Figure 5.10.	Differential <i>PLA2G5</i> mRNA Expression Associated with rs525380 Measured in all ASAP Study Tissues	224
Figure 5.11.	The Per Allele Effects of rs525380 on <i>PLA2G5</i> mRNA Expression in Mammary Artery, Aorta Media and Aorta Adventitia	225
Figure 5.12.	Relative Differential Expression of <i>PLA2G2A</i> and <i>PLA2G5</i> mRNA with rs11573185 and rs525380 in the ASAP Study	226
Figure 5.13.	The Per Allele Effects of rs11573185 and rs525380 on <i>PLA2G2A</i> mRNA in the Liver	227
Figure 5.14.	Meta-Analysis Results for the Association of rs11573185 and LDL	229
Figure 5.15.	Meta-Analysis Results for the Association of rs11573185 and HDL	229
Figure 5.16.	Meta-Analysis Results for the Association of rs11573185 and Log(n) Triglycerides	230
Figure 5.17.	Meta-Analysis Results for the Association of rs11573185 and CHD	230
Figure 5.18.	Meta-Analysis Results for the Association of rs11573185 and Log(n) sPLA2 Mass	231
Figure 5.19.	Meta-Analysis Results for the Association of rs11573185 and Log(n) sPLA2 Activity	231
Figure 5.20.	Meta-Analysis Results for the Association of rs525380 and CHD	232
Figure 5.21.	Meta-Analysis Results for the Association of rs525380 and Log(n) sPLA2 Mass	233
Figure 5.22.	Meta-Analysis Results for the Association of rs525380 and Log(n) sPLA2 Activity	233
Figure 6.1A.	The Gene Dense Region Including <i>PLA2G10</i> on Chromosome 16 Protein Coding Transcript	240
Figure 6.1B	The Gene Dense Region Including <i>PLA2G10</i> on Chromosome 16 Protein Nonsense Mediated Decay	240
Figure 6.2.	Meta-Analysis Results for the Association of rs4003228 and Log(n) sPLA2-IIA Mass	244
Figure 6.3.	Meta-Analysis Results for the Association of rs4003228 and Log(n) sPLA2 Activity	245

Figure 6.4.	Meta-Analysis Results for the Association of rs4003228 and CHD	246
Figure 6.5.	<i>PLA2G10</i> mRNA Expression Associated with SNPs Measured Across ASAP Study Tissues	247
Figure 7.1.	Comparison of <i>PLA2</i> mRNA Expression Profiles in ASAP Study	258

List of Tables

Table 1.1.	An Overview of Known and Predicted CHD Risk Factors	38
Table 1.2	Structural Features and Enzymatic Properties of sPLA2 Enzymes	54
Table 1.3	The Diverse Nature of sPLA2 Isoenzyme activity in Disease	64
Table 1.4.	Studies Investigating sPLA2-IIA Mass/sPLA2 Activity associations with CHD	67
Table 1.5.	Previously Identified <i>PLA2G5</i> Tagging-SNPs	85
Table 2.1.	Commonly Used Reagents/Stocks/Products/Kits	91
Table 2.2.	Individual TaqMan Genotyping Assays	96
Table 2.3.	Individual TaqMan Gene Expression Assays	96
Table 2.4.	Commonly Used Websites/Software/Services	96
Table 2.5.	Commonly Used Laboratory Machines and Equipment	98
Table 2.6.	The Studies Included in sPLA2-IIA Related Meta Analyses	136
Table 2.7.	Circulating Biomarkers/Traits Measured in Each Study	137
Table 2.8.	Studies Included in sPLA2-V Related Meta-Analyses	140
Table 2.9.	Studies Included in sPLA2-X related Meta-Analyses	141
Table 3.1A.	Results for rs11573156 C with MatInspector	149
Table 3.1B.	Results for rs11573156 G with MatInspector	149
Table 3.2.	Results from PITA Predicting miRNA Binding in the Region of rs11573156	154
Table 3.3.	Results from miRTar Predicting miRNA Binding in the Region of rs11573156	155
Table 3.4.	Results from RegRNA Predicting miRNA Binding in the Region of rs11573156	156
Table 3.5a	Results for rs3767221 A with MatInspector	160
Table 3.5b	Results for rs3767221 C with MatInspector	160
Table 3.6.	Results from miRTar Predicting miRNA Binding in the Region of rs3767221	163

Table 3.7.	Results from PITA Predicting miRNA Binding in the Region of rs3767221	164
Table 4.1.	Results from PRISMA Guided Analysis for Varespladib	187
Table 4.2.	<i>PLA2G2A</i> Functional SNPs Associated with sPLA2-IIA Mass and sPLA2 Activity	189
Table 4.3.	Baseline Measurements for EPIC-Norfolk and GRACE-France	192
Table 4.4A	The Intermediate Trait Measurements for EPIC-Norfolk by rs11573156 and rs3767221 Genotype	193
Table 4.4B	The Intermediate Trait Measurements for GRACE-France by rs11573156 and rs3767221 Genotype	194
Table 4.5A	The Association Between Tertiles of sPLA2 Activity and Traditional CVD Risk Factors	201
Table 4.5B	The Association Between Tertiles of sPLA2-IIA Mass and Traditional CVD Risk Factors	202
Table 5.1.	Univariate Analysis for <i>PLA2G5</i> Tagging-SNPs in UDACS	214
Table 5.2.	A Small Initial Meta-Analysis of <i>PLA2G5</i> Tagging SNPs	228
Table 6.1.	Association of <i>PLA2G10</i> SNPs with sPLA2 Activity in EPIC-Norfolk and GRACE-France	241
Table 6.2.	Association of <i>PLA2G10</i> SNPs with sPLA2 Mass in EPIC-Norfolk and GRACE-France	242
Table 6.3.	Associations between <i>PLA2G10</i> Tagging-SNPs and CHD in EPIC-Norfolk	242

All reprinted Figures and Tables are reprinted with permission of the original publisher.

List of Acronyms

AA	Arachidonic Acid
ABCA1	ATP-binding cassette transporter
AMC-PAS	Academic Medical Centre Amsterdam Premature Atherosclerosis Study
ANOVA	Analysis of Variance
Apo	Apolipoprotein
ASAP	Advanced Study of Aortic Pathology
BHF-FHS	British Heart Foundation Family Heart Study
Bp	Base Pair
BPI	Bactericidal Permeability-Increasing Protein
BRHS	British Regional Heart Study
BWHHS	British Women's Health and Heart Study
cAMP	Cyclic adenosine monophosphate
CAD	Coronary Artery Disease
cDNA	Complementary DNA
CCHS	Copenhagen City Heart Study
CD36	Cluster of Differentiation 36 (Class B Scavenger Receptor)
CETP	Cholesteryl Ester Transfer Protein
CHD	Coronary Heart Disease
CL	Cardiolipins
CM	Chylomicron
CNV	Copy Number Variation
CRP	C-Reactive Protein
CVD	Cardiovascular Disease
DNA	Deoxyribose Nucleic Acid
EAS	Edinburgh Artery Study

EPIC	European Prospective Investigation into Cancer
EST	Expressed Sequence Tag
FAST-MI	French Registry on Acute ST-elevation Myocardial Infarction
GENDEMIP	GENetic DEtermi- nation of Myocardial Infarction in Prague
GENDER	GENetic Determinants of Restenosis
GRACE	Global Registry of Acute Coronary Events
GRAPHIC	Genetic Regulation of Arterial Pressure of Humans in the Community
GWAS	Genome Wide Association Study
HDL	High Density Lipoprotein
IHCS	Intermountain Heart Collaborative Study
IDL	Intermediate Density Lipoprotein
IL	Interleukin
IMPROVE	IMPROVE Study
iPLA2	Calcium Independent Secretory Phospholipase A2
KAROLA	Langzeiterfolge der KARdiOLOGischen Anschlussheilbehandlung
kDa	Kilodaltons
LCAT	Lecithin-Cholesterol Acyltransferase
LDL	Low Density Lipoprotein
LDLR	Low Density Lipoprotein Receptor
LIFE-Heart	Leipzig Heart Study
LP(a)	Lipoprotein (a)
LPL	Lipoprotein Lipase
LpPLA2	Lipoprotein-associated phospholipase A2
M	Molar
MADGE	Microplate Array Diagonal Gel Electrophoresis
Medstar	Med-Star Study
MERLIN-TIMI 36	Metabolic Efficiency With Ranolazine for Less Ischemia in

	Non-ST-Elevation Acute Coronary Syndromes
MI	Myocardial Infarction
miRNA	micro-RNA
mM	Millimolar
MR	Mendelian Randomisation
mRNA	Messenger RNA
MVE	Major Vascular Event
NANSAID	Non-Aspirin Non-Steroidal Anti-Inflammatory Drugs
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NEFA	Non-Esterified Free Fatty Acid
nM	Nanomolar
NPHSII	Northwick Park Heart Study
NO	Nitric Oxide
NOS	Nitric Oxide Synthase
NTC	No Template Control
ox-LDL	Oxidized Low Density Lipoprotein
PAI-1	Plasminogen Activator Inhibitor-1
PAF	Platelet Activating Factor
PennCath	University of Pennsylvania Catheterization study program
PC	Phosphatidylcholine
PCR	Polymerase Chain Reaction
PE	Phosphatidylethanolamine
PI	Phosphatidylinositol
PLA2	Phospholipase A2
PREVEND	Prevention of Renal and Vascular End-stage Disease Intervention Trial
PROSPER	PROspective Study of Pravastatin in the Elderly at Risk
PROVE-IT TIMI 22	Pravastatin Or atorVastatin Evaluation and Infection Trial

PS	Pulmonary Stenosis
RAAS	Renin–Angiotensin–Aldosterone System
Rotterdam	Rotterdam Study
RNA	Ribonucleic acid
RCT	Randomised Control Trial
ROS	Reactive Oxygen Species
SD	Standard Deviation
SEM	Standard Error Mean
sd-LDL	Small-Dense Low Density Lipoprotein
SMART	Second Manifestations of ARTerial disease
SNP	Single Nucleotide Polymorphism
sPLA2	Secretory Phospholipase A2
SR-BI	Scavenger receptor class B Type 1
T2D	Type Two Diabetes
T-Cells	T-Lymphocytes
TG	Triglycerides
TNF	Tumour Necrosis Factor
tpa	Tissue plasminogen activator
TPT	Thrombosis Prevention Trial
T-SNP	Tagging Single Nucleotide Polymorphism
UCP Study	Utrecht Coronary Prevention Study
UDACS	University College London Diabetes And Cardiovascular Disease Study
UTR	Untranslated Region
VLDL	Very Low Density Lipoprotein
WHHL	Watanabe heritable hyperlipidemic
Whitehall II	Whitehall Phase II Study
WHO	World Health Organisation

Publications Resulting From This Thesis

Accepted for Publication

Exeter, HJ., L. Folkersen, et al. (2012). "Functional Analysis of Two *PLA2G2A* Variants with Secretory Phospholipase A2-IIA Levels" PLoS ONE
<http://dx.plos.org/10.1371/journal.pone.0041139>

Submitted- Under Review

Holmes M.V., T. Simon, H.J.Exeter, et al. (2012) "The Role of Secretory Phospholipase A2 in Coronary Heart Disease" Submitted December 2012.

CHAPTER 1

INTRODUCTION

1. Cardiovascular disease: A Worldwide Problem

Cardiovascular disease is an umbrella term for all diseases of the heart and vascular system, including coronary heart disease (CHD) and stroke. It is a well reported and supported fact that cardiovascular diseases are currently the number one cause of mortality globally, since CVDs are now one of the biggest medical problems in low income (developing economy) countries (Figure 1.1) as well as in high income (developed economy) countries (Figure 1.2) (WHO; Media Centre, <http://www.who.int/mediacentre/factsheets/fs317/en/>). In 2008 a World Health Organisation (WHO) study showed that 17.3 million deaths worldwide could be attributed to CVDs, and a third of these were specifically identified as being caused by CHD (WHO; Media Centre, <http://www.who.int/mediacentre/factsheets/fs317/en/>).

In order to reduce the high mortality rates attributed to CHD it is important to consider all that we currently know about the disease, its prevention and its treatment and then to attempt to find novel contributing factors that could be used to improve on this current knowledge. This chapter will introduce what is already known about CHD and its known risk factors and then discuss what is known about the novel risk factor, secretory phospholipase A2. The purpose of the rest of the thesis is to investigate what is currently unknown about secretory phospholipase activity, expression and the true causality of its association with CHD.

1.1. Coronary Heart Disease and Atherosclerosis

Coronary heart disease (also known as coronary artery disease) is defined by a stenosis of the vessels that supply blood and oxygen to the heart, leading to tissue necrosis and cardiac events such as myocardial infarction. Stenosis is usually caused by an accumulation of plaque in the arteries. This is commonly known as atherosclerosis. When atherosclerosis occurs in the coronary arteries it can lead to CHD and is the underlying cause of CHD in most instances. The plaque is mostly composed of fatty material, leukocytes and other cell debris (which is explained in more detail in **1.2.2**) and the build-up of plaque causes the walls of the coronary arteries to stiffen and narrow, preventing blood from flowing efficiently (Weber and Noels 2011).

World Health Organisation 10 Leading Causes of Death (2008) Low Income Countries

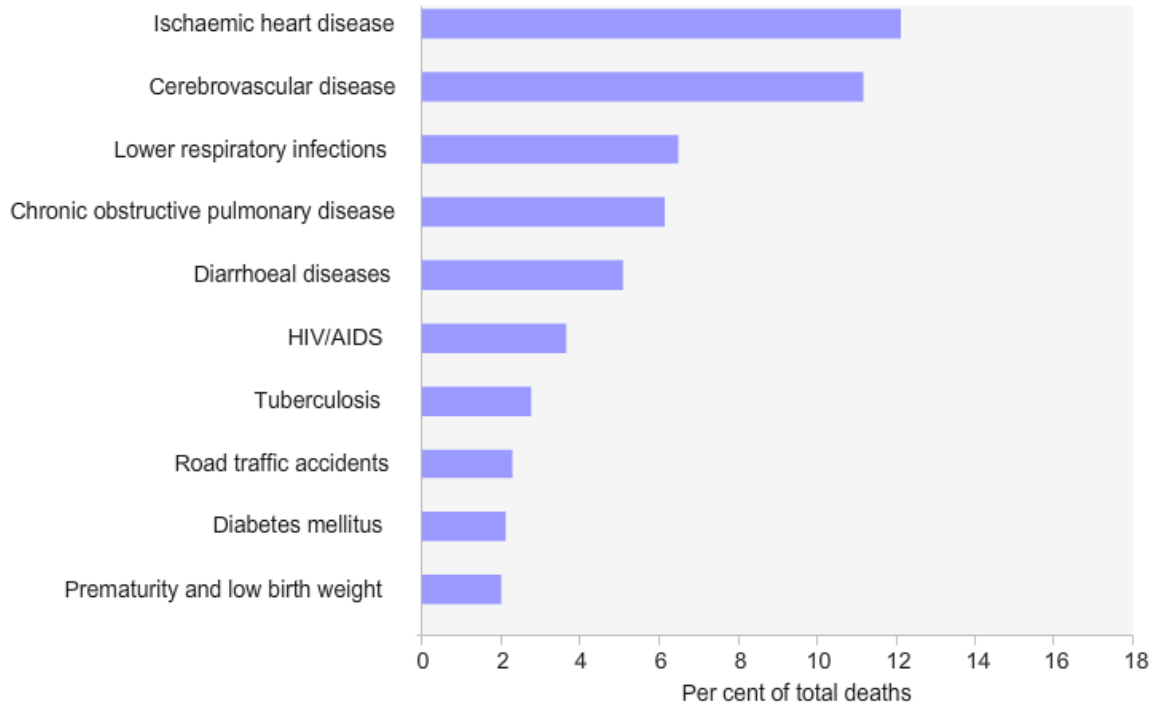


Figure 1.1. Graphic to show the top 10 causes of mortality in low income countries in 2008. Graphic and data according to the World Health Organisation- Heart disease is the biggest cause of death (www.who.int) reprinted with permission.

World Health Organisation 10 Leading Causes of Death (2008) High Income Countries

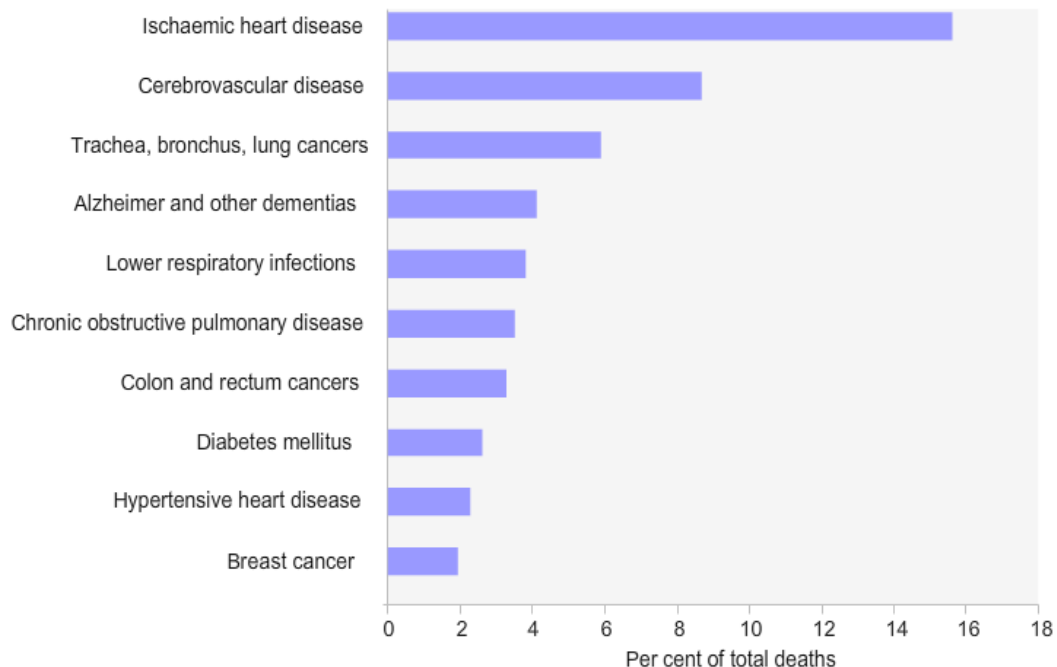


Figure 1.2. Graphic to show the top 10 causes of mortality in high income countries in 2008. Graphic and data according to the World Health Organisation- Heart disease is the biggest cause of death (www.who.int) reprinted with permission.

1.2. Pathology of Atherosclerosis

1.2.1. The Morphology of a Normal Artery

Atherosclerosis is a disease of the arteries. Therefore in order to understand the development of an atherosclerotic plaque it is important to first examine the structure of a normal artery. A normal artery is composed of several distinct layers surrounding the lumen, each incorporating different cell types (Figure 1.3.). The inner layer of the artery is the intima. The intima is separated from the lumen by a thin layer of endothelial cells known as the endothelium and on the peripheral boundary by a sheet of elastic fibres known as the internal elastic lamina. The intima is a relatively thin layer (the normal intima size has been exaggerated for better visualisation in Figure 1.3) comprised of extracellular matrix connective tissues, a composition largely made up of collagen fibres and proteoglycans, which often have a role in the trapping and stability of proteins within the layer. The middle layer, or media, is composed entirely of smooth muscle cells, which maintain local blood pressure and control the volume of the vessel. This layer is thicker in arteries than it is in veins as arteries are carrying blood pumped away from the heart towards the organs and must therefore be able to stand the higher pressure. The outer layer of the normal artery is the adventitia. This is a mixture of connective tissues, fibroblasts and smooth muscle cells (Lusis 2000).

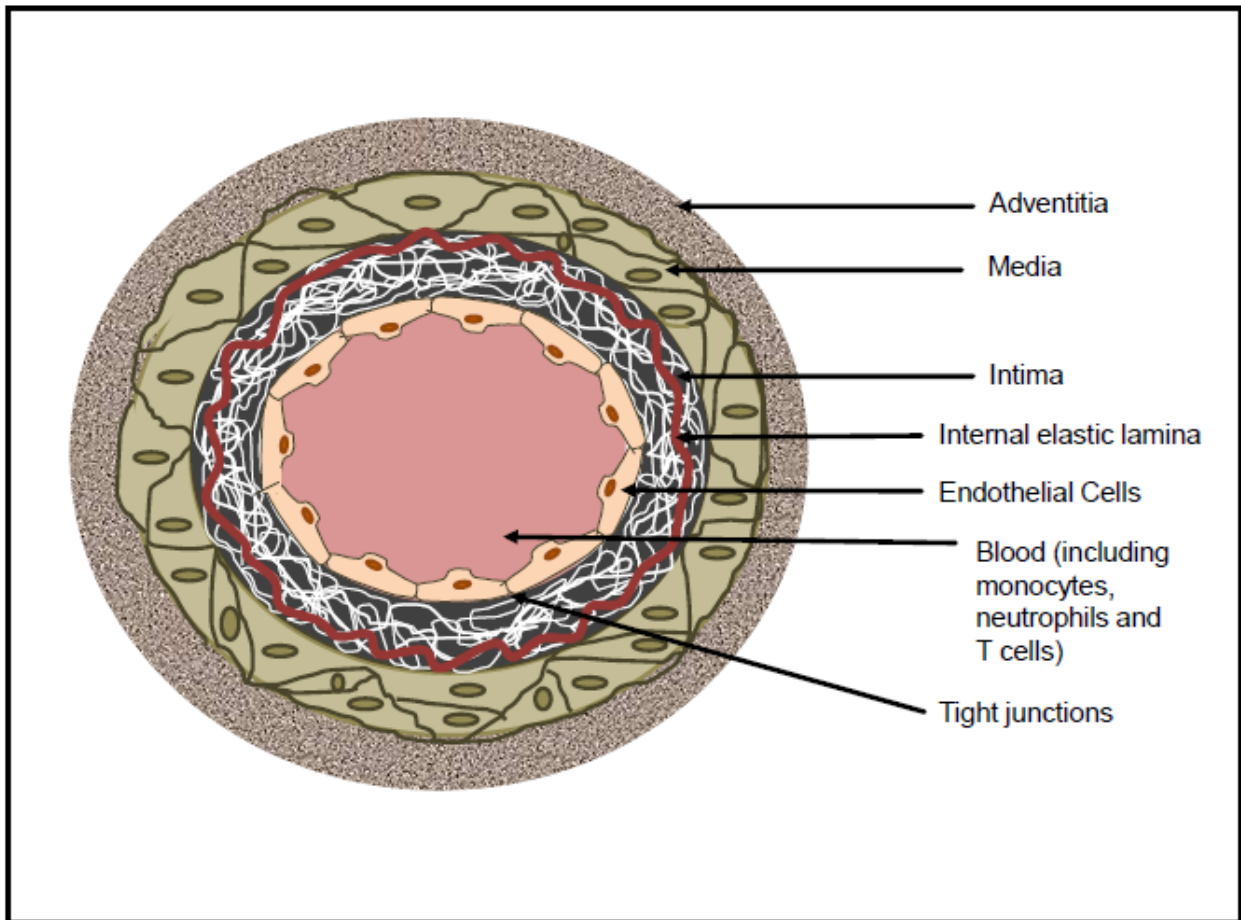


Figure 1.3. The morphology of the normal artery (Lusis 2000). The size of the intima has been exaggerated for better visualization. The adventitia is the outermost layer composed of connective tissue, fibroblasts and smooth muscle cells. The middle layer is the media and is made up of smooth muscle cells. The media contributes to blood pressure regulation. The inner layer is the intima, which contains an internal elastic lamina and is lined at the innermost boundary by a thin layer of epithelial cells, which are joined by tight junctions, which help to regulate what passes from the lumen into the arterial wall and *vice versa*. Blood flows through the centre of the vessel in the lumen.

1.2.2. Development of Atherosclerosis

Atherosclerosis is a multifactorial disease that progresses over multiple stages and is defined by an accumulation of lipid and fibrous matter in the large arteries. Onset of atherosclerosis is triggered by several complex signal cascades including the inflammatory pathway, endothelial function, platelet activation, lipid metabolism and oxidative stress. These pathways operate independently of one another, but also have the ability to ‘cross-talk,’ providing positive feedback for each other and driving plaque formation (Davies and Woolf 1993; Tegos, Kalodiki et al. 2001; Mudau, Genis et al. 2012). These pathways and their regulatory mechanisms are discussed below. The family of secretory phospholipase A2s are thought to increase the risk of atherosclerosis as part of an inflammatory cascade as

well as by contributing to the disruption of the lipoprotein metabolism pathway (**1.5.7**), though it is debatable whether these actions are causal for atherogenesis or as a result of it.

The Step-Wise Progression of an Atherosclerotic Plaque

There are four main stages of atherosclerotic plaque progression; Initiation, LDL modification and accumulation, foam cell formation and necrotic core formation leading to plaque rupture. These stages are described in detail below (Ross 1999).

Plaque Initiation by Endothelial Dysfunction

The endothelium plays a major role in the regulation of vascular homeostasis. The endothelium has several important functions including balancing vasoconstriction and vasodilation and the homeostasis of smooth muscle cell proliferation and migration. It also plays a role in fibrinolysis and thrombogenesis (Cannon 1998; Mudau, Genis et al. 2012).

Endothelium dysfunction usually refers to an interruption in the synthesis or release of nitric oxide (NO) which is constitutively synthesised by the endothelium and has a wide range of biological properties that make it the ideal mechanism by which the endothelium can control both vascular homeostasis and anti-thrombosis of platelets (Cannon 1998; Mudau, Genis et al. 2012). Disruption of any of the endothelial functions can trigger processes that are directly related to an increase in atherosclerosis. Thus, disruption to the endothelium is likely to induce the onset of atherosclerosis (Sitia, Tomasoni et al. 2010).

There are several chemical and mechanical changes within the blood flow in the vessel that can contribute to atherosclerosis development by endothelial dysfunction (**1.3**) (Sitia, Tomasoni et al. 2010), including an increase in low density lipoprotein (LDL) cholesterol, LDL modifications, hypertension, diabetes, the introduction of free radicals through cigarette smoking, genetic variations, increased homocysteine concentrations and infection. All of these potential causes have the ability to elicit an inflammatory response and the 'injury' to the endothelium promotes atherogenesis since the endothelium becomes more adhesive for leukocytes and platelets, more permeable and produces further growth factors and cytokines, which are known procoagulant molecules and important regulators of inflammation (Luscher and Barton 1997; Ross 1999; Sitia, Tomasoni et al. 2010).

Since endothelial dysfunction tends to induce an inflammatory response and the onset of atherosclerosis, atherosclerosis is considered to be an inflammatory disease (Libby 1995; Hansson 2009) and this is apparent at the initiation of plaque development. Increased inflammatory mediators are produced by endothelial injury as well as increased platelet and leukocyte aggregation. Platelets secrete chemokines and cytokines, which are known

inflammatory mediators and can be activated by secretions from the vascular wall, creating feedback loops that can interfere with normal vascular regulation (Lusis 2000; Langer and Gawaz 2008; Zhang 2008). This inflammatory activation of endothelial cells leads to proatherogenic cells and proinflammatory mediators such as intracellular and vascular cell adhesion molecules, IL1- β , TNF- α and the acute phase protein C-reactive protein (CRP) to be recruited and secreted (Lusis 2000; Langer and Gawaz 2008; Zhang 2008). Additional inflammatory mediators, such as sPLA2, may also be involved in the regulation of these signal cascades. If prolonged, this inflammatory activation of endothelial cells (leading to proatherogenic cells and proinflammatory mediators) will initiate the development of cholesterol-rich macrophages, known as foam cells, which is a primary stage of plaque formation (Lusis 2000; Langer and Gawaz 2008; Zhang 2008) and continues throughout the development of the plaque. The earliest atherosclerotic lesions can be seen to develop in humans within the first decade of life. At this point the lesions are known as 'fatty streaks' and are not yet of clinical relevance. These early fatty streaks are distinguished by high smooth muscle cell proliferation and migration towards the intima as well as increased collagen synthesis (Figure 1.4A) and they appear subendothelially (Zhang 2008).

Development of the Plaque: LDL Accumulation and Modification

Low density lipoprotein (LDL) is the major transporter of cholesterol in humans and delivers cholesterol to the peripheral cells of the body (Steinberg 1987). In circulation, LDL has the ability to diffuse passively through endothelial cell tight junctions into the intima. Normally these LDL particles are metabolised and do not to accumulate in the vessel wall (Brown and Goldstein 1986; Brown and Goldstein 1997; Lusis, Mar et al. 2004). However, when there is an unusually high level of LDL in the circulation or endothelial dysfunction increases LDL permeability, both transportation and retention of LDL in the vessel wall seem to increase (Lusis 2000). LDL retention in the intima is further increased by LDL modifications including oxidation (Berliner and Heinecke 1996; Badimon, Storey et al. 2011). The increase in modified LDL and inflammatory mediators attracts monocytes to first adhere to the endothelium in these areas and then traverse the vessel wall to enter the intima (Figure 1.4B) (Pentikainen, Oorni et al. 2000; Hansson 2009).

Foam Cell Formation

Once monocytes are induced to enter the vessel wall by the increase in modified LDL and inflammatory mediators they then differentiate into macrophages, which form foam cells by cellular uptake of atherogenic lipoproteins. These lipoproteins are classified as different chemical modifications of ApoB containing lipoproteins, such as lipoprotein(a), small dense

LDL (sd-LDL), oxidised LDL (ox-LDL) and enzymatically modified LDL (Orso, Grandl et al. 2011).

Macrophages are classified as foam cells when they are engorged with lipid droplets, which has been shown experimentally to be due to enzyme modified LDL loading of macrophages and, additionally, ox-LDL loading of macrophages. This forms lipid laden foam cells and these foam cells play an important role in the synthesis and progression of atherosclerotic plaques (Orso, Grandl et al. 2011). Secretory phospholipases have been shown previously to interact with LDL to produce enzyme modified LDL (sd-LDL) (Hurt-Camejo, Camejo et al. 2001). Some secretory phospholipases have also been shown to have an increased affinity for ox-LDL binding, resulting in further modifications to LDL particles (Wootton, Arora et al. 2007; Blache, Gautier et al. 2012). Therefore secretory phospholipases appear to be strong candidates to play a role in atherosclerosis progression (1.5.7).

Plaque progression: Necrotic Core Formation and Plaque Rupture

Initially the vessel copes with the formation of an atherosclerotic plaque by remodelling itself. Arteries can expand outwards while maintaining the size of the lumen (Figure 1.4C). However this is a finite process and eventually the plaque will grow big enough to encroach upon the lumen and block circulation (Libby 1995).

As atherosclerotic plaques increase in size they form fibrous caps (Ross 1999; Businaro, Tagliani et al. 2012), separating the plaque from the lumen. The larger plaques induce growth factor release from leukocytes and platelets (Ross 1999; Businaro, Tagliani et al. 2012), which stimulate smooth muscle cell proliferation and these cells form a fibrous cap over the plaque. The core of the lesion now includes foam cells, leukocytes and other cell debris- caused by a combination of necrosis and apoptosis (Ross 1999; Businaro, Tagliani et al. 2012). Mature plaques can be considered stable and unstable (Ross 1999; Businaro, Tagliani et al. 2012). Stable plaques are identified by their thick fibrous caps, making it unlikely that a rupture will occur and spill atherosclerotic matter into the vessel. Unstable plaques have thinner fibrous caps and are more prone to rupture. Plaque rupture releases tissue fragments including clot-promoting and platelet activating factors (Figure 1.4D). Post-rupture thrombosis ensues and clots further narrow the vessel, reducing the flow of oxygenated blood able to flow through the vessel, thus increasing the likelihood of cardiac tissue death and a cardiac event. Additionally, each plaque develops its own micro-vessel network and haemorrhaging and clotting of these micro-vessels may contribute to vessel narrowing (Borissoff, Spronk et al. 2011). The process of atherosclerotic plaque formation to thrombus is shown below (Figure 1.4 a, b, c and d, taken from (Ross 1999)).

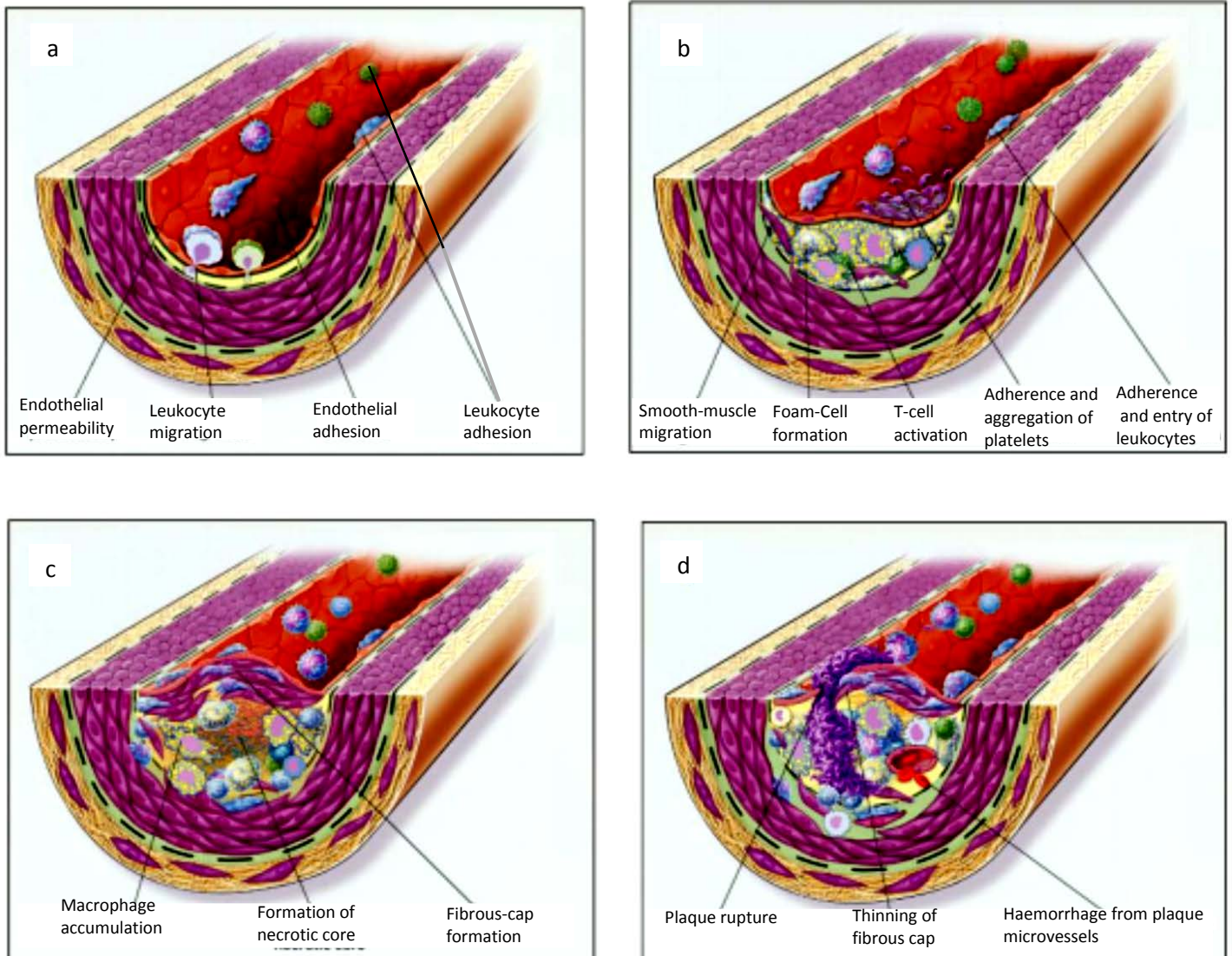


Figure 1.4. Atherosclerotic Plaque Progression (Reproduced with permission from Ross et al. 1999, NEJM, copyright Massachusetts Medical Society). a) Leukocyte and monocyte cells migrate from the lumen into the intima, forming a fatty streak. b) Foam cell formation by the uptake of LDL cholesterol/modified-LDL/ox-LDL by macrophages, as well as smooth muscle cell stimulation and migration, T cell activation, platelet aggregation and the adherence and migration of leukocytes cause the lesion to increase in size and protrude into the vessel. c) Apoptosis and necrosis promote the formation of the necrotic core. Foam cells continue to form and the proliferation of smooth muscle cells forms the fibrous cap. d) If the fibrous cap thins the plaque will rupture inducing platelet activating factor and a thrombus forms, occluding the vessel and preventing blood flow. Secretory phospholipases are inflammatory enzymes and may contribute to plaque progression by contributing to the induction of leukocyte migration (a) and by foam cell formation by the modification of LDL particles (b).

1.3. Predicting the Risk of CHD

1.3.1. Some known and Emerging Risk Factors of CHD

CHD is a complex multivariate disease and a number of elements combine to trigger its onset. These elements are known as CHD risk factors and the concept of predicting the risk of CHD in an individual has evolved from numerous epidemiological studies of these factors over the past 50 years or so. One of the best known epidemiological studies is the Framingham Heart Study. In the 1960s the Framingham Heart Study showed several risk factors were key independent predictors of CHD and these included age, cigarette smoking, hypertension, increased LDL-cholesterol and decreased high density lipoprotein (HDL) cholesterol. Independent risk factors in combination provide a comprehensive risk profile for each individual. In the decades that followed the first Framingham results many other independent CHD risk factors have been established (Table 1.1) and subsequent studies have also identified a number of emerging risk factors that may have the potential to independently affect CHD risk. For example, a recent study used an in-depth proteomics platform on 800 women with CHD and 800 women with stroke who are enrolled in the Women's Health Institute study. The outcome was the identification of the protein B2M as a potential risk marker for CHD (Prentice, Paczesny et al. 2010).

Over the years there have also been many suspected risk factors that have been proved not to be independently contributing to the onset of CHD. These factors tend to be markers of CHD or a known independent risk factor for CHD rather than being directly causal (Graham 2006).

Known causal risk factors can be genetic or environmental and both types interact to increase the risk of CHD. Independent risk factors and their pathways are excellent targets for the treatment and prevention of CHD. However, current available treatments have not reduced CHD mortality rates significantly enough and it is therefore necessary to continue investigating other contributing risk factors.

Risk Factor (Biological)	Confirmed?	Evidence
Environment Related		
Diet	✓	F.B. Hu, International Encyclopaedia of Public Health; pages 181-190
Smoking	✓	Leif Erhardt. Atherosclerosis: 2009; 205, p23-32
Lack of Exercise/Obesity	✓	(Duscha, Slentz et al. 2005)
Gender	✓	Lloyd-Jones, DM; Larson, MG et al. LANCET: 1999; 353(9147) 89-92
Lipids:		
Low density lipoprotein	✓	Brown & Goldstein, Science 1986, Martin et al. Med Clin North Am, 1996(1): 13-26
High density lipoprotein	✓ (Inverse)	Lewington S, Whitlock G, Clarke R, et al. Lancet. 2007; 370: 1829–1839
Triglycerides		Triglyceride Coronary Disease Genetics, Emerging Risk Factors et al. Lancet 2010; 375(9726): 1634-1639
Haemostasis/Thrombosis:		
Fibrinogen	✗	Reinhart WH. Vascular Medicine 2003; 8(3), 211-216
von Willebrand Factor		
Tissue plasminogen activator		
D-Dimer		
Plasminogen activator inhibitor 1	✓	Bart De Taeye, L Harris Smith, Douglas E Vaughan. Current Opinions in Pharmacology., 2005; 5(2) 149-154
Oxidative Stress/Others:		
Homocysteine	✓	Veeranna V, Zalawadiya SK, Niraj A et al. J Am Coll Cardiol. 2011 Aug 30;58(10):1025-33.
Lipoprotein-associated phospholipase A2	✗	The Lp-PLA ₂ Studies Collaboration, Lancet. 2010 May 1; 375(9725): 1536–1544.
Myeloperoxidase		
Family History	✓	Hawe E, Talmud PJ, Miller GJ, Humphries SE; Ann Hum Genet. 2003 Mar;67(Pt 2):97-106.
Hypertension	✓	Donald M. Lloyd-Jones, Eric P. Leip, et al. Hypertension. 2005; 45: 39-45
Diabetes and Glucose	✓	Garcia et al. 1974. Park et al. Laboratory Investigation; 2012: 92: 1171-1180
Age	✓	Jousilahti P, Vartiainen E et al. Circulation: 1999; 99:1165-1172
Inflammation:		
Interleukins		
C-Reactive Protein	✗	C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC), BMJ 2011; 342:d548
Soluble Adhesion Molecules		
Secretory Phospholipase A2*		

***Secretory phospholipase A2 is the factor under investigation for this thesis.**

Table 1.1. An overview of some of the known and predicted risk factors for CHD. If evidence has proved causality or not for a risk factor it has been indicated with a tick for 'causal' or a cross for 'non-causal' followed by a reference for evidence in each case. Those risk factors with no tick or cross are still currently under investigation.

1.3.2. Environmental Risk Factors

An example of an environmental risk factor for the progression of atherosclerosis is diet, since this is the source of exogenous cholesterol (**1.4.5**). In excess this cholesterol will cause a build-up of foam cells. Cholesterol, modified cholesterol particles and inflammatory cells, such as macrophages induce foam cell formation by accumulating in the artery along with calcium and fibrous tissue. Increased triglycerides in the diet can also exaggerate these plaque forming effects (Bhupathiraju and Tucker 2011).

Cigarette smoking is another highly publicised and well researched environmental risk factor for CHD as well as ischemic stroke, many cancers and other vascular diseases (Doll and Hill 1966; Ockene and Miller 1997). Smoking is strongly associated to risk of CHD and is one of the single most modifiable factors that, if ceased, can significantly reduce CHD risk (Gordon, Kannel et al. 1974). There are more than 250 by-products of smoking and these can affect CHD risk in different ways. For example, products of tobacco smoke are known to lead to increased endothelial dysfunction and activation of adhesion molecules, required for monocyte and platelet attachment (Lu and Creager 2004). Damage to the lining of the vessels caused by smoking increases the risk of atherosclerosis by narrowing the lumen. Some of the potential pathways thought to be mediated by cigarette smoking include; the deposition of free radicals directly from the smoke and the activation of endogenous sources of free radicals, such as uncoupled nitrous oxide synthase (NOS) and NADPH oxidase, which increases oxidative stress. Cigarette smoke also induces the activation of inflammatory mediators such as neutrophils, monocytes, platelets and T-cells, which in turn causes an increase in cytokines. Overall these interactions can cause inflammatory mediator activation (which may include inducing secretory phospholipase expression, though this theory would require experimental validation), vasomotor dysfunction, increased pro-thrombotic and decreased fibrinolytic factors, an increase in leukocyte and platelet activation, increased lipid peroxidation and smooth muscle cell proliferation (Ambrose and Barua 2004).

1.3.3. Factors with a Strong Genetic Component.

Due to the complex nature of CHD and its underlying cause (atherosclerosis), many different biological processes are involved. Considering this, it is likely that in different individuals and in different populations, variations within the genes coding for the multiple implicated mechanisms contributing to CHD will have an impact on CHD risk. Evidence for this is seen in large scale investigations into the many 'candidate genes' and genetic loci identified as being involved in processes related to atherosclerosis (Yang 2012). These include genes related to inflammation (Humphries, Luong et al. 2001), lipid metabolism

(Lusis 2012) and thrombosis (Kathiresan, Yang et al. 2006). The volume of data produced about these associations is vast and too much to summarise completely here. However the potential role of genetics in atherosclerosis is discussed below.

A family history of CHD is a strong independent risk factor for future CHD events (Hawe, Talmud et al. 2003) in both men and women (Boer, Feskens et al. 1999), especially in the case of early onset CHD (Pohjola-Sintonen, Rissanen et al. 1998). This has been supported by evidence from several observational studies (Pohjola-Sintonen, Rissanen et al. 1998; Boer, Feskens et al. 1999; Hawe, Talmud et al. 2003) as well as twin studies comparing CHD risk between sets of monozygotic and dizygotic twins. Many heritable genetic factors have been identified and studied (Willer, Sanna et al. 2008) and more potential genetic traits are currently being discovered and investigated, such as those of the *PLA2* genes, which are explored in this thesis. The sum of these genetic risks will give a strong indication of CHD susceptibility. Family history reflects the combination of heritable genetic factors and environmental factors. Environmental factors can enhance or suppress the effects of genetic variants dependent on lifestyle choices and socio-economic status.

The independent association of cardiovascular disease with hypertension has been known for some time and a meta-analysis of over a million people and 61 prospective studies confirms this association (Lewington, Clarke et al. 2002). The mechanisms by which hypertension increases the risk of these cardiovascular diseases include decreased NO, increased oxidative stress and changes in the renin angiotensin aldosterone system (RAAS) (Mason 2011). These pathways are regulated by gene interactions and this has been verified in gene studies. There have been several genetic variations identified as being involved in the RAAS system that have been studied in conjunction with their effects on CHD risk and show strong, independent associations (Jones, Dhamrait et al. 2003).

Genetic studies like these help to identify potential pathways to stimulate or inhibit in order to prevent or treat CHD and similar research could be applied to novel biomarkers such as secretory phospholipase A2s.

Inflammatory Markers of Atherosclerosis

Since atherosclerosis is considered an inflammatory disorder, the inflammatory pathway plays a crucial role in atherogenesis and atherosclerotic plaque maintenance by the recruitment of inflammatory cells and mediators, such as monocyte- derived macrophages and T-lymphocytes, into the arterial intima (Figures 1.4B and 1.4C). Both epidemiological and clinical studies have validated a strong association between inflammatory markers related to the inflammatory cascades known to induce and maintain atherogenesis and CHD

events (Kuller and Meilahn 1996), (Willerson and Ridker 2004). Inflammatory markers are also potentially involved in disease progression and secondary events. Many inflammatory markers have been the subject of recent studies to identify associations with CHD events, for example, interleukins 6 and 18 (IL-6 and IL-18) the soluble CD-40 ligand, CRP, lipoprotein-associated phospholipase A2 (Lp-PLA2) and secretory phospholipase A2 (sPLA2) (Corson 2009). Genetic studies of several of these and additional inflammatory markers have identified some polymorphisms associated with CHD (Andreotti, Porto et al. 2002). Secretory phospholipase A2 is the inflammatory marker that is the focus of the following chapters of this thesis (**1.5**).

Despite the association between inflammation and atherosclerosis, the evidence for anti-inflammatories having a positive effect on coronary heart disease risk is mixed. For example, a study of the anti-inflammatory properties of statins has shown that their anti-inflammatory effect seems to contribute to the overall CHD risk reduction that is reported (Jain and Ridker 2005) (although the majority of this effect is attributed to the lipid lowering effect of statins) (Jain and Ridker 2005). Contrary to this, an observational study of non-aspirin, non-steroidal anti-inflammatory drugs (NANSAID) provided evidence to suggest that there is no cardio-protective effect provided by these drugs and they are therefore not good candidates to aid the prevention of CHD (Ray, Stein et al. 2002). This evidence is partly contradicted by a prospective case-control NANSAID study, which shows a reduction in the risk of MI for patients using NANSAIDs without aspirin. However this study further concludes NANSAIDs offer no further cardio-protection to patients already taking aspirin and recommends that further studies should be carried out (Kimmel, Berlin et al. 2004). Overall, the evidence from these studies suggests that new anti-inflammatory compounds need to be assessed on their individual properties and associations with atherosclerosis/CHD as well as in combination with known drugs to assess their true benefits.

Diabetes and Glucose Control as a CHD Risk Factor

Diabetes has a strong genetic component and there has been a steady interest in the study of the positive association between diabetes and coronary heart disease for some time (Ali, Narayan et al. 2010). The pre-dominant form of diabetes is Type 2 diabetes (T2D). Results from the Framingham Heart Study showed strong evidence that individuals with diabetes have a two to four fold increase in risk of coronary heart disease when compared with non-diabetics (Kannel and McGee 1979). Further to this, results are presented from a study by Moss et al. that show risk of mortality from cardiovascular diseases is higher in diabetics compared to a non-diabetic group (Moss, Klein et al. 1991). A first incident of a cardiovascular event, such as myocardial infarction (MI), is known to significantly increase

the risk of a second event, with the 10 year risk of CHD in a man increasing from 3.8% to approximately 20% (Pekkanen, Linn et al. 1990; Rosengren, Hagman et al. 1997). A study of patients with T2D with no previous coronary heart disease revealed that their risk of CHD was as high as a comparative group of non-diabetics with a previous cardiovascular event (Haffner, Lehto et al. 1998).

This strong association between CHD mortality rates and diabetes may indicate that the two diseases share common precursors; excess inflammation, insulin resistance and visceral adiposity and mechanistic processes such as renal dysfunction and oxidative stress are thought to be common to both diseases (Ali, Narayan et al. 2010). This may also mean that there is some cross-over between the genetic regulation of these diseases too (Singh, Singh et al. 2007). There has been some speculation that secretory phospholipases may play a role in Type II diabetes onset and/or progression and, if so, whether this association is a result of a causal role in atherosclerosis or due to a direct association with diabetes. These ideas require experimental verification (Rosenson and Hurt-Camejo 2012).

1.4. Lipoprotein Metabolism

A significant body of evidence confirms that lipid abnormalities are central to the development of CHD and its associated disorders (Lusis 2000; Pentikainen, Oorni et al. 2000). Lipid metabolism plays a key role in lipid regulation and therefore CHD initiation and progression (Davies and Woolf 1993; Aulchenko, Ripatti et al. 2009).

Lipids are required in the body for cell membrane formation, hormone synthesis and bile acid production and the major classes of lipid are; cholesterol, triglycerides (TG) and phospholipids. Due to the hydrophobic nature of lipids they are not soluble in blood and must be packaged into lipid/protein complexes known as lipoproteins to allow them to travel efficiently in the circulation. (Nichols 1969; Lin, Mousa et al. 2010).

1.4.1. Structure of Lipoproteins

Figure 1.5 illustrates a typical lipoprotein. The major common lipoprotein complexes are chylomicrons, very low density lipoproteins (VLDL), intermediate density lipoproteins (IDL), LDL and HDL. (Nichols 1969; Lin, Mousa et al. 2010). These five major classes of lipoprotein are based on size and density and include several subclasses. The largest class are the chylomicrons (CM), which are synthesised in the intestine and are approximately 75-1200nm in diameter and have a density of less than 0.95g/ml. The second largest class is very low density lipoprotein (VLDL), which are synthesised in the liver with a mean diameter of 30-80nm and a density of 0.95-1.006g/ml. These two classes of lipoprotein are classified

as triglyceride rich and the catabolism of either group synthesises another class of lipoprotein, chylomicron remnants or intermediate density lipoprotein (IDL), which have a mean diameter of 25-35nm and a density of 1.006-1.019g/ml (Nichols 1969; Freedman, Otvos et al. 1998; Carmena, Duriez et al. 2004). Low density lipoprotein (LDL) is the next largest lipoprotein with a mean diameter of 18-25nm and is a product of VLDL catabolism in the liver. There are subclasses of LDL with different densities. These include; LDL I, which is 1.02-1.03g/ml, LDL II is 1.03-1.04g/ml and LDL III, which is 1.04-1.06g/ml (Freedman, Otvos et al. 1998; Carmena, Duriez et al. 2004). The smallest lipoprotein is high density lipoprotein (HDL), which is synthesised in the liver and intestine. These particles have a mean diameter of 5-12nm and are split subclasses including, HDL 2, which is 1.063-1.125g/ml and HDL 3, which is 1.125-1.210g/ml (Freedman, Otvos et al. 1998; Carmena, Duriez et al. 2004). Together LDL and HDL are classified as cholesterol ester and phospholipid-rich lipoproteins (Nichols 1969).

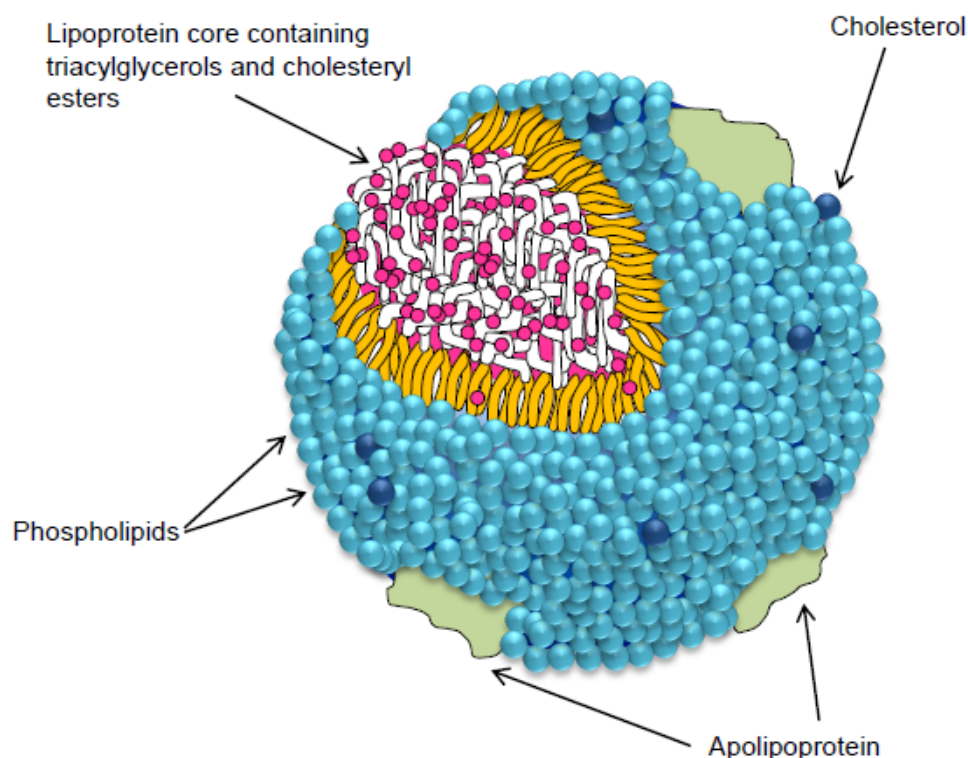


Figure 1.5. A circulating plasma lipoprotein. Lipoprotein particles are composed of phospholipids, free cholesterol, cholesteryl esters, TG and proteins from a class called apolipoproteins, which includes Apo A-I, A-II, A-IV, A-V, B₄₈, B₁₀₀, C-I, C-II, C-III and E (see **1.4.2.**) Each lipoprotein contains core triglycerides and cholesteryl esters. A phospholipid surface layer with interspersed cholesterol and one or more apolipoproteins.

1.4.2. Phospholipids

The major groups of phospholipids are; glycerophospholipids and sphingolipids (van Meer, Voelker et al. 2008). Glycerophospholipids are considered the most abundant and are present in virtually all mammalian cell membranes (van Meer, Voelker et al. 2008). There are at least 10 different structures that are defined by the structure of their head groups (Hermansson, Hokynar et al. 2011). The major classes in mammalian cells are; phosphatidylcholine (PC), phosphatidylethanolamines (PE), phosphatidylinositols (PI), phosphatidylserines (PS) and cardiolipins (CL).

1.4.3. Metabolism of Lipoproteins

Three major pathways work to control lipoprotein metabolism (Figure 1.6). Each pathway works independently, but all three are connected in such a way that interference with one upsets the homeostasis and will ultimately affect the other pathways too. The three pathways are: the transport of dietary (exogenous) fat, the transport of hepatic (endogenous) fat and reverse cholesterol transport (Kwiterovich 2000).

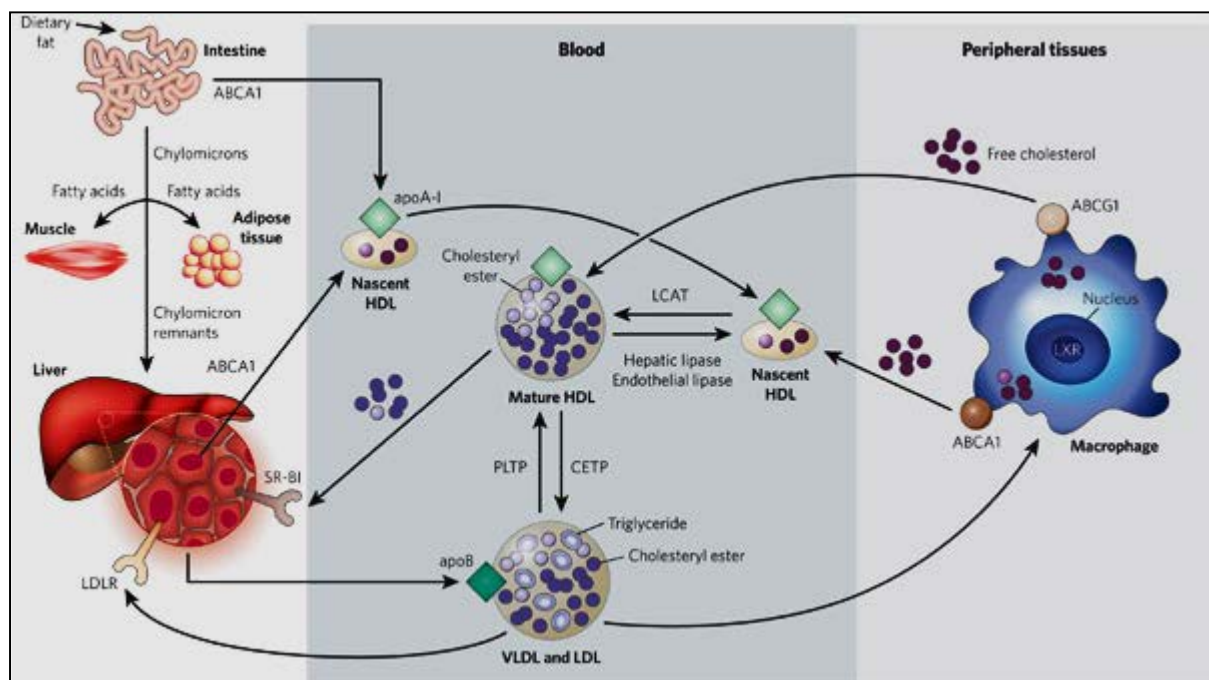


Figure 1.6. Lipoprotein Metabolism. An overview of the lipoprotein metabolism process illustrating both the endogenous (formed in the liver) and exogenous (dietary intake) pathways of lipid/lipoprotein synthesis as well as the removal of lipoproteins from circulation back to the liver for processing, by the reverse cholesterol transport pathway. ABCA1 = ATP-binding cassette transporter. ABCG1 = ATP-binding cassette transporter sub-family G. SR-B1= scavenger receptor B1. PLTP= phospholipase transporter protein. CETP= cholesterol ester transfer protein. LDL= low density lipoprotein. HDL= high density lipoprotein. LDLR= LDL receptor. LCAT = Lecithin-cholesterol acyltransferase. Schematic reprinted with permission from (Rader and Daugherty 2008).

Transport of Dietary Fat

Dietary fat is absorbed by the intestine, where cholesterol and fatty acids are re-esterified to form triglycerides and cholesteryl esters. These are packaged with ApoB-48, phospholipids, unesterified cholesterol and several further apolipoproteins into nascent chylomicrons (CM). These are transported to the Golgi apparatus through vesicles and next released to the lymph for transportation into the blood circulation. Once in the circulation the CMs exchange apolipoproteins with HDL including; ApoE, ApoC-I, ApoC-II, ApoC-III and ApoA-V. The enzyme lipoprotein lipase (LPL) hydrolyses the core triglycerides into free fatty acids and glycerol when activated by ApoC-II (Eisenberg and Levy 1975; Eisenberg 1984). The CM is then broken down into CM remnants (Kwiterovich 2000; Kindel, Lee et al. 2010) and these CM remnants are taken up via the LDL receptor (LDLR) and the LDLR related protein with ApoE as the ligand. The majority are taken up in this way by the liver parenchymal cells (Brown and Goldstein 1986; Beisiegel, Weber et al. 1989; Krieger and Herz 1994). Interruption to this pathway leading to a reduction in CM clearance may result in elevated TG levels (post-prandial) and this is a recognized feature of CAD patients (Patsch, Miesenbock et al. 1992; Yu and Cooper 2001).

Transport of Endogenous Fat

A process that occurs in the liver is the synthesis and secretion of VLDL particles and this process incorporates ApoB-100 rather than ApoB-48. VLDL is triglyceride rich and once secreted into the plasma the triglyceride in VLDL is catabolised by the lipoprotein lipase enzyme and its ApoC-II co-factor into free fatty acids and glycerol. The breakdown of the TG produces smaller VLDL remnants and eventually IDL. IDL particles are taken back up into the liver via the interaction of ApoE and LDLR. Further hydrolysis of triglycerides by the hepatic lipase enzyme leads to the production of LDL and small dense LDL particles. The LDL particles are removed from the circulation either by cells requiring cholesterol or by the liver when there is an excess, by the interaction of ApoB-100 with the LDLR. However, if LDL is oxidised it can also enter the macrophage through the scavenger receptors CD36 and SR-A (Kwiterovich 2000).

Reverse Cholesterol Transport and the Role of HDL

The phrase 'reverse cholesterol transport' (RCT) was first coined by Glomset in 1968 (Glomset 1968) and describes the process by which cholesterol is transported from non-hepatic, peripheral cells, including macrophages, back to the liver for processing and excretion in the bile and faeces. HDL is the lipoprotein associated with this process. It acquires free cholesterol from the peripheral, non-hepatic tissues by either passive diffusion

or interaction with either the SR-BI receptor or the ABCA1 transporter. The ABCA1 transporter system is considered the most efficient pathway. It is responsible for around half the cholesterol efflux from macrophages to HDL (Moreno, Sanz et al. 2009). HDL particles are the densest due to their high protein:cholesterol ratio and their most abundant proteins are ApoA-I and ApoA-II. When newly secreted HDL is poorly lipidated and acquires free cholesterol from the peripheral cells via ABCA1. The free cholesterol is then esterified by lecithin-cholesterol acyltransferase (LCAT), with ApoA-I as a co-factor (Fielding and Fielding 1995). Mature HDL cholesterol particles also transfer esterified cholesterol to other ApoB containing lipoproteins, in exchange for TGs, by the cholesteryl ester transporter protein (CETP), which can then be taken up by the liver via the LDL or LRP receptors and increases the efficiency of the RCT system. By removing cholesterol from peripheral tissues (especially macrophages) in this manner, HDL and the RCT system have been shown to have an inverse relationship to CAD risk and increased HDL cholesterol has been shown to reduce the risk of CVD (Moreno, Sanz et al. 2009). Despite the seemingly positive research showing low HDL to be a potentially causal risk factor for CVD, recent research has challenged the concept that raising plasma HDL will lower the risk of myocardial infarction (Voight, Peloso et al. 2012).

1.4.4. Lipoprotein Metabolism as a CHD Risk Factor

Since an early step in atherogenesis is the accumulation of lipoprotein particles and their aggregates in the arterial intima, lipoprotein metabolism is a key process in the initiation and progress of atherosclerosis and several of the most significant CHD risk factors are associated with this process (Lusis, Mar et al. 2004). Generally, increased LDL/modified LDL (predominantly sd-LDL particles) and triglycerides in combination with reduced HDL are thought to be representative of an atherosclerotic lipid profile. However, despite increased TG being cited as part of this atherosclerotic risk profile, it is not clear whether it is independently associated with CHD, or merely a marker for the disease (Sarwar, Danesh et al. 2007; Kawakami, Osaka et al. 2008; Triglyceride Coronary Disease Genetics, Emerging Risk Factors et al. 2010; Libby, Ridker et al. 2011).

There are many enzymes, cellular receptors and ligands involved in the metabolism and clearance of lipoprotein particles (Packard, Caslake et al. 2000) and over the last few years studies have identified genes associated with lipoproteins, lipoprotein metabolism and lipoprotein modification to be excellent 'candidate genes' with the potential to alter the balance of lipid metabolism and, therefore, CHD risk (Willer, Sanna et al. 2008; Aulchenko, Ripatti et al. 2009; Interleukin-6 Receptor Mendelian Randomisation Analysis, Hingorani et al. 2012). One group of genes associated with lipoprotein modification and metabolism are

the genes coding for the secretory phospholipase A2 enzymes. Secretory phospholipase A2 enzymes play a significant role in the modification of lipoproteins and the speed at which they are cleared from the vessel. Their role in lipid metabolism has led to the question of what effect these enzymes may have on the risk of CHD (**1.5.7**).

1.5. The Secretory Phospholipase A2 Family

Having described the process of atherosclerosis and some of the associated risk factors above, a good candidate to investigate as a novel CHD risk factor has emerged as the sPLA2 enzyme family. These enzymes are involved in both the inflammatory pathway and in lipoprotein metabolism as described in detail below.

Secretory phospholipase A2 (sPLA2) is a structurally heterogeneous family of proteins belonging to the PLA2 superfamily. This superfamily comprises 15 groups also including; cystolic, calcium independent (iPLA2) and platelet-activating (PAF) acetyl hydrolase/oxidized lipid lipoprotein associated phospholipase A2 (LpPLA2) (Burke and Dennis 2009). There are approximately 30 PLA2 enzymes split between these groups and around a third of these are considered sPLA2 enzymes (Murakami, Taketomi et al. 2011). Regarding cardiovascular disease, both iPLA2 and cystolic PLA2 are thought to contribute indirectly through the production of various lipid mediators in different cell types (Mallat, Lambeau et al. 2010). However, a more direct role has been attributed to both sPLA2 and Lp-PLA2 due to the additional ability of these enzymes to bind to and hydrolyse the phospholipid moieties of lipoproteins (Mallat, Lambeau et al. 2010). Secretory PLA2 enzymes have the ability to bind and hydrolyse the phospholipid components of both unmodified lipoproteins and oxidized-LDL (ox-LDL), whereas LpPLA2 specifically hydrolyses ox-LDL and does not affect unmodified lipoproteins (Mallat, Lambeau et al. 2010).

The sPLA2 enzymes are secreted by various mammalian cells and tissues in a lipid-free state. The sPLA2 family is identified by the common ability to catalyse the hydrolysis of the middle (sn-2) ester bond of glycerophospholipids (Figure 1.7) in lipoproteins and in some cases, cell membranes, producing non-esterified free fatty acids (NEFAs) and lysophospholipids, which are pre-cursors of pro-inflammatory lipid mediators, including arachidonic acid derived eicosanoids such as prostaglandins and leukotrienes (Six and Dennis 2000; Lambeau and Gelb 2008) (Figure 1.8). They can also act as secondary messengers in intracellular inflammatory pathways (Hurt-Camejo, Camejo et al. 2000). For example it has been suggested by analysis of data from animal studies that sPLA2s including IIA, V and X could stimulate inflammatory pathways by binding to murine, specific membrane receptors (including the M-type receptor, which is expressed by neutrophils, monocytes and macrophages) and that both normal and catalytically inactive sPLA2s can

induce the production of cytokines and chemokines, inducing further inflammation (Mallat, Lambeau et al. 2010). Secretory PLA2s are secreted and require variable μM -order Ca^{2+} which binds a common calcium peptide loop and interacts with surrounding water molecules for catalysis (Lambeau and Gelb 2008). It is therefore reasonable to suggest that the primary target phospholipids for sPLA2 are extracellular.

Secretory phospholipases can act through either an autocrine or a paracrine pathway (Jaulmes, Janvier et al. 2005; Murakami, Taketomi et al. 2011) to release arachidonic acid, an important pre-cursor of lipid mediators, or a variety of saturated, monounsaturated and polyunsaturated fatty acids such as eicosapentaenoic acid and docosahexaenoic acid (Murakami, Taketomi et al. 2011). These products are, conversely to arachidonic acid, pre-cursors of anti-inflammatory lipid mediators, suggesting that there is a small anti-inflammatory element to sPLA2 actions in addition to the overwhelming evidence for the overall pro-inflammatory status of sPLA2. The accompanying lysophospholipid product may be, for example, lysophosphatidylcholine or lysophosphatidic acid which have various bioactivities (Murakami, Taketomi et al. 2011).

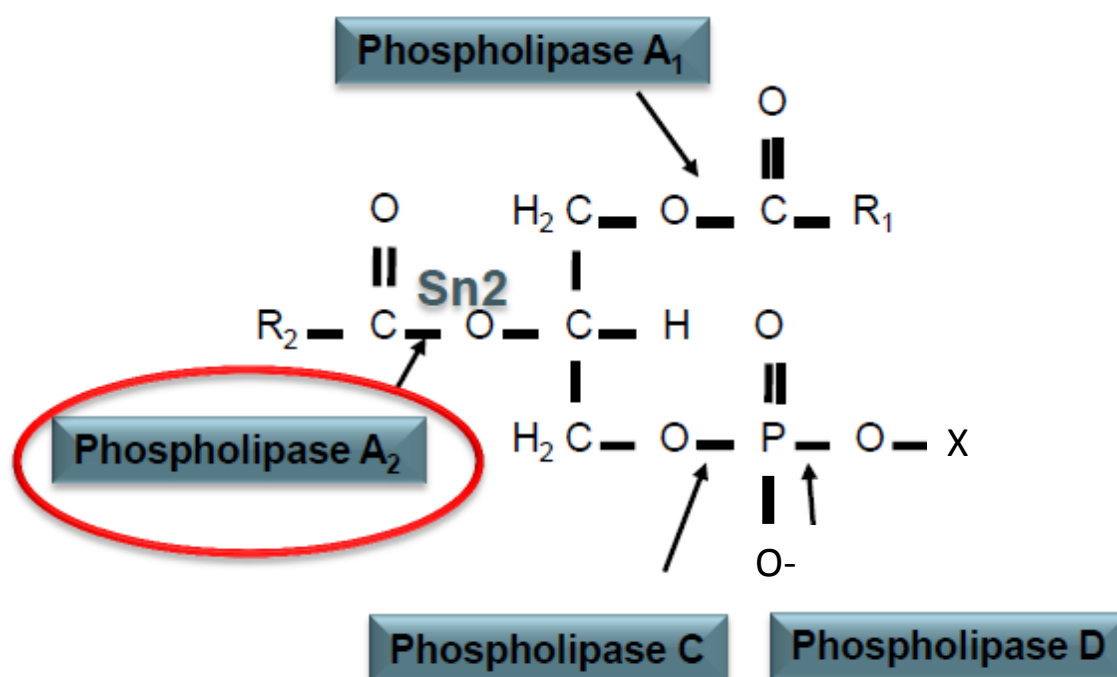


Figure 1.7. The site of action for sPLA2. Secretory phospholipase A2 acts to catalyse the hydrolysis of the Sn-2 ester bond of phospholipids. This reaction produces a lysophospholipid and a free fatty acid. Hydrolysis of this particular bond distinguishes PLA2 from the other phospholipases.

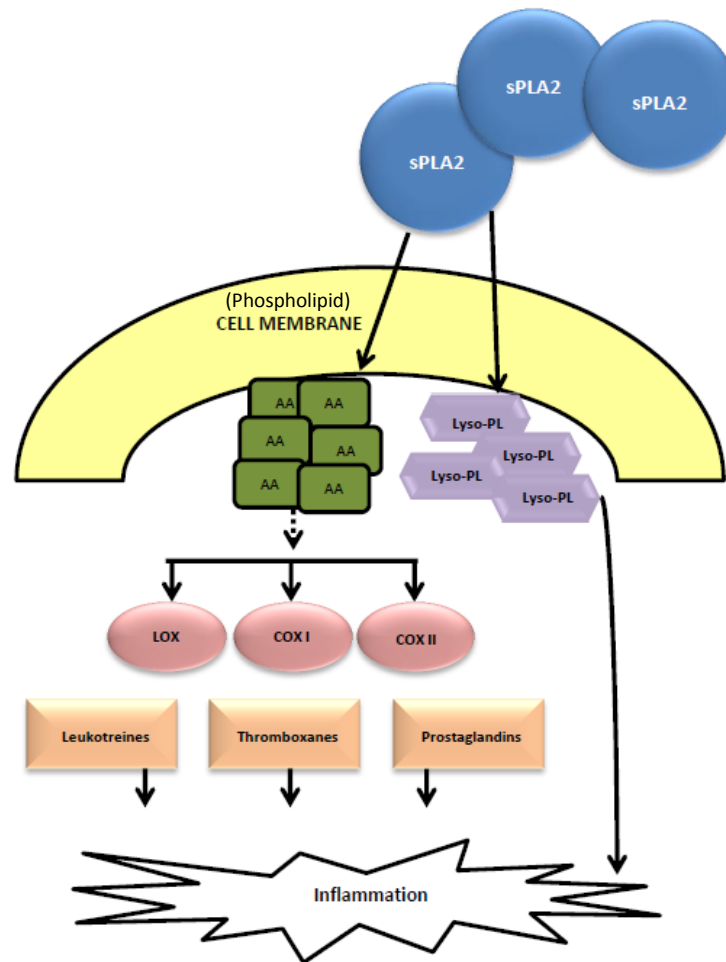


Figure 1.8. This figure represents the potential outcomes of sPLA2 hydrolysis of cell surface phospholipids. The enzyme acts on phospholipids to produce lysophospholipids (Lyso-PL) and non-esterified free fatty acids (NEFA) such as arachidonic acid (AA). As well as these products having independent proinflammatory activity, AA is a precursor of pro-inflammatory lipid mediators such as the AA-derived eicosanoids; leukotrienes, thromboxanes and prostaglandins, all of which lead to increased inflammation.

The secretory phospholipases are also disulphide-rich enzymes with molecular weights ranging between ~14 and 16kDa, with the exception of sPLA2-III, which has a molecular weight of ~55kDa. Human sPLA2 proteins are generally well conserved, and share homologs with mouse sPLA2 proteins (Lambeau and Gelb 2008). Secretory PLA2s can be induced by inflammatory signals, such as tumour necrosis factor (TNF) and interleukin-1 (Il-1) (Hurt-Camejo, Camejo et al. 2001; Jaulmes, Thierry et al. 2006) and have been shown to be inhibited by growth factors such as transforming growth factor β (Vervoordeldonk, Schalkwijk et al. 1996). To date, 11 sPLA2s have been identified in

mammals; IB, IIA, IIC, IID, IIE, IIF, III, V, X, XIIA and XIIB (Murakami, Taketomi et al. 2011). The oldest sPLA2 isoenzyme appears to be IB, whose orthologues are found in *C.elegans* and then, slightly later, sPLA2-X which appeared in amphibians (Lambeau and Gelb 2008). Both of the DNA structures for these sPLA2 groups have encoded an N-terminal propeptide, but at later stages evolutionary pressure saw the emergence of sPLA2 group II, which were encoding sPLA2 enzymes without an N-terminal propeptide. This group first appeared in snakes. The most recently formed isoform is sPLA2-V, which is also considered the simplest isoform and was found in mammals (Murakami, Taketomi et al. 2011). The secretory PLA2 groups are shown below in Figure 1.9.

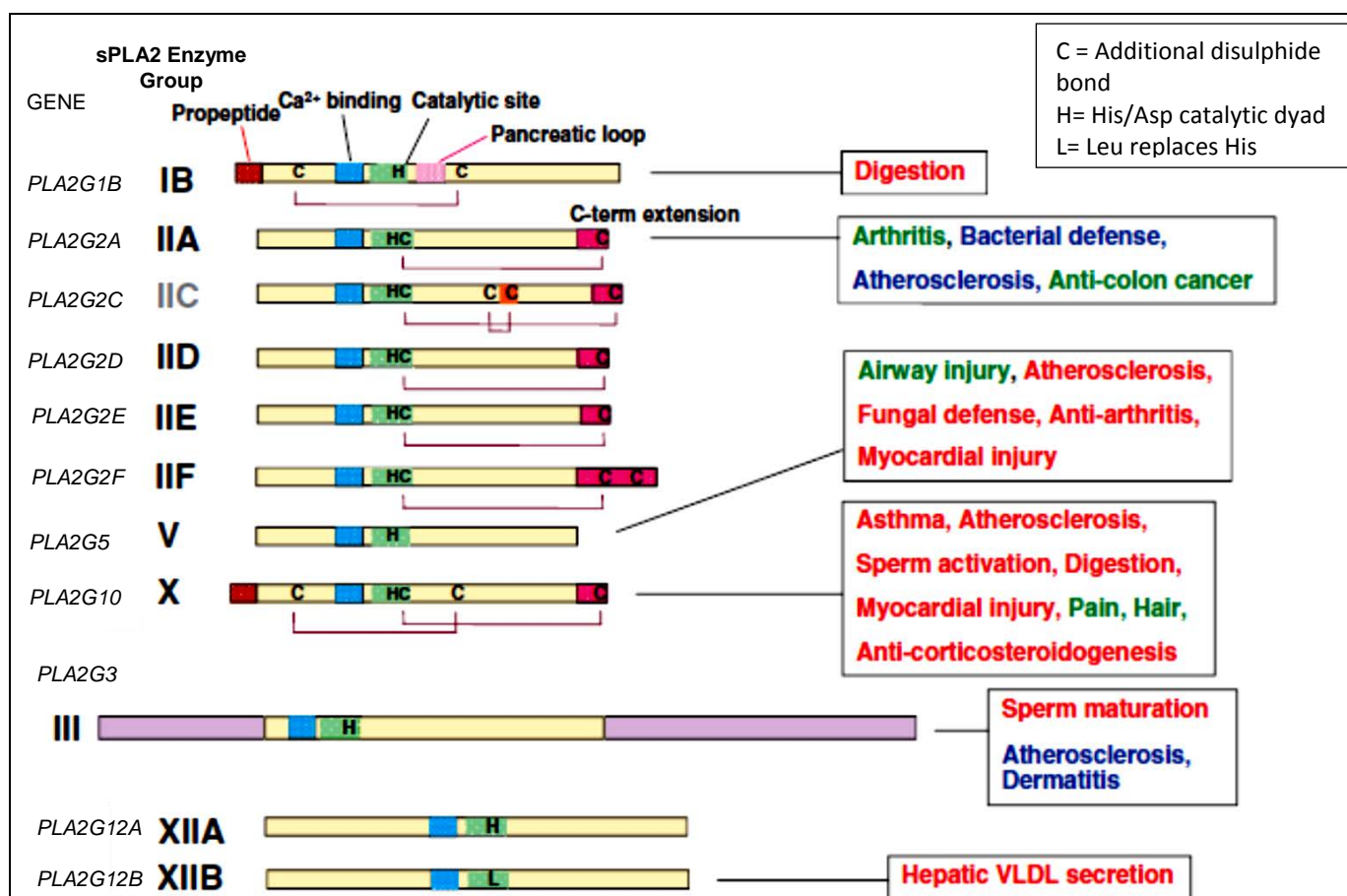


Figure 1.9. A schematic of the different secretory phospholipase A2 isoforms and their roles in the body and in disease; assessed by knockouts (red), transgenic models (blue) or both (green). All enzymes contain a conserved catalytic histidine-aspartate dyad and a Ca²⁺ binding loop. Group IB contains a unique pancreatic loop, which is thought to be a redundant structure conserved by this group. Group II is characterized by the absence of an N-terminal propeptide (which exists in groups IB and X) and the presence of cysteine49-cysteine within the C terminal extension. Group IIC has a unique insertion with an extra disulphide bond and is absent in humans. sPLA2-V is most similar to Group II, but lacks the group II-specific disulphide and C-terminal extension. sPLA2-X has both group I and II properties. This schematic is reprinted with permission from (Murakami, Taketomi et al. 2011)).

Human groups IIA, V and X share a common three dimensional structure and the catalytic mechanism of sPLA2 involves several key features. The common structure of the active site includes an Aspartate-Histidine catalytic dyad (Figure 1.10), unique to the sPLA2 family, and the previously described peptide loop bound Ca²⁺ ion (Lambeau and Gelb 2008). Secretory PLA2 enzymes are water soluble, unlike their preferred substrates, phospholipids, whose hydrophobic tails prevent them from water dissolution (unless in a conformation such as a lipoprotein). In order for phospholipids to become water soluble, a second, interfacial binding site has been determined by previous experimental analysis in addition to the active site which allows the enzyme to bind to the intracellular surface of the phospholipid substrate, significantly improving the binding ability of the substrate to the active site

(Lambeau and Gelb 2008). Differences in the binding abilities of human sPLA2 groups to different phospholipids occur due to their diverse side-chains of sPLA2 enzymes, which create distinctive charges associated with each sPLA2 group. These differences in charge dictate the ability of each group to interact with anionic and/or zwitterionic substrates. For example, phosphatidylcholine is zwitterionic whereas phosphatidylglycerol and phosphatidylserine are anionic (Lambeau and Gelb 2008).

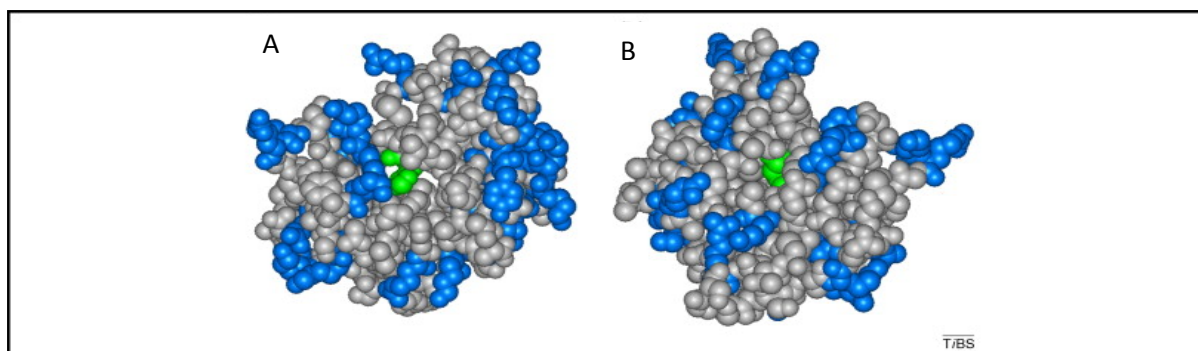


Figure 1.10. This figure shows the Aspartate-Histidine catalytic dyad active site of secretory phospholipase A2 Group IIA in green. This catalytic dyad is specific to the sPLA2 family of enzymes. The figure shows the interfacial binding surface from the front (A) and back (B). Lysine (cationic residues) are shown in blue and the crystal structure is taken from (Wery, Schevitz et al. 1991; Birts, Barton et al. 2010). The ability of sPLA2 enzymes to interact with different lipoproteins effectively depends on differences in enzyme conformation. For example the sPLA2-IIA has less tryptophan residues than sPLA2-V or sPLA2-X and tryptophan has been shown to increase the binding ability of sPLA2 to PC rich phospholipids. Reprinted with permission from (Lambeau and Gelb 2008).

A previous study (Sawada, Murakami et al. 1999) showed that there are different profiles of sPLA2 isoenzyme expression and one form of the enzyme may partially compensate for the loss/lack of another (Sawada, Murakami et al. 1999). However, there are catabolic differences between each isoenzyme regarding phospholipid hydrolysis, and therefore their interactions with lipoproteins (Lambeau and Gelb 2008).

Secretory PLA2-IIA shows the weakest hydrolysing effect towards the surface phospholipids of lipoproteins comparatively. This has previously been attributed to the fact that wild type sPLA2-IIA has less ability to hydrolyse the major lipoprotein membrane phospholipid, PC compared to sPLA2-V and X. The presence of increased tryptophan residues occurring naturally in sPLA2-V and X compared to sPLA2-IIA has been cited as a reason why there is a difference in the hydrolysis between these sPLA2s (Hurt-Camejo, Camejo et al. 2001; Bezzine, Bollinger et al. 2002). The suggested mechanism by which

sPLA2-V and X can hydrolyse PC is by preferentially binding these tryptophan residues to the interfacial region of PC-rich phospholipid bilayers (Pruzanski, Lambeau et al. 2005). A previous study examining this relationship used sPLA2-IIA proteins modified with additional tryptophan residues. Results showed sPLA2-IIA proteins modified with extra tryptophan residues at position 3 or 31 had an increased ability to hydrolyse PC molecules. Significantly the tryptophan residues at these positions improve sPLA2-IIA hydrolysis of PC-rich membranes in a way that is highly comparable with the action of sPLA2-V (Beers, Buckland et al. 2003). It is unknown how much of each sPLA2 enzyme is bound to lipoprotein phospholipids at any one time. The binding abilities of each of the sPLA2 enzymes, groups IIA, V and X, to common phospholipids is shown below (Table 1.2).

	Structural Features			Enzymatic Properties				
	Molecular Mass, kDa	Active Site	PE Hydrolysis	PC Hydrolysis	PAF Hydrolysis	oxPL Hydrolysis	Proteoglycan Binding	Foam Cell Formation
sPLA2-IIA	13.9	His/Asp	+++	+/-	-	?†	+++	+
sPLA2-V	13.8	His/Asp	+++	+++	-	?†	++	++
sPLA2-X	13.6	His/Asp	+++	+++	++	?†	-	++

PE = phosphatidylethanolamine; PC = phosphatidylcholine, PAF=platelet activating factor

Phosphatidylcholine and phosphatidylethanolamine are differentially hydrolysed by sPLA2s.

sPLA2-X has a unique capacity to hydrolyse PAF.

On the basis of studies with pancreatic and snake venom sPLA2s, their activity on oxidized phosphatidylcholine may be weaker than on nonoxidized phosphatidylcholine.

sPLA2-IIA and -V, but not -X, show high affinity for heparan sulphate proteoglycans.

LDLs modified by sPLA2-IIA, -V, and -X are efficiently internalized by macrophages to induce the accumulation of cellular cholesterol ester and formation of foam cells.

?† Detailed analysis of the enzymatic properties of human sPLA2s toward ox-PLs has not been performed.

Table 1.2. Shows the structural features and enzymatic properties of secretory phospholipase A2 groups IIA, V and X. The three enzymes share a similar molecular weight and an identical His/Asp active site. However, each enzyme has a differential binding ability to different substrates. This is probably due to the subtle structural differences between them, such as additional tryptophan residues on sPLA2-V and X and the improved interfacial binding of these enzymes, which increases the ability of these enzymes to bind to zwitterionic phospholipids in comparison to sPLA2-IIA. The high proteoglycan binding ability of sPLA2-IIA and, to a lesser extent sPLA2-V improves the binding ability of these enzymes to common lipoprotein surface phospholipids. Increased potency of sPLA2-V and X for PC and PE (compared to sPLA2-IIA) leads to increased LDL modifications and a higher risk of foam cell formation (from Mallat et al. 2010).

In this thesis the enzymes focused on are; sPLA2-IIA, sPLA2-V and sPLA2-X as there is evidence to suggest an association between these enzymes and atherosclerosis (1.5.7).

1.5.1. Secretory phospholipase A2 Group IIA (sPLA2-IIA)

The common amino acid sequence for the sPLA2-IIA protein is shown in Figure 1.11. Secretory PLA2-IIA is probably the most researched and understood of the sPLA2 enzymes. It is expressed in and secreted by cells of numerous mammalian tissues including the ovaries, heart, liver and prostate and has been identified in many cell types, including

macrophages (Kimura-Matsumoto, Ishikawa et al. 2008), as well as extracellularly in the circulation (Menschikowski, Hagelgans et al. 2006). The initial isolation of sPLA2-IIA was from human rheumatoid arthritis synovial fluid (Wery, Schevitz et al. 1991; Masuda, Murakami et al. 2005), then later in the serum of patients with severe inflammatory disorders such as sepsis and septic shock (Pruzanski, Vadas et al. 1985; Waydhas, Nast-Kolb et al. 1989). The correlation of enzyme levels with the severity of the patients' inflammatory disorder suggests that sPLA2-IIA has a mediating role to play in the inflammatory pathway by responding to inflammatory cytokines and lipolysaccharides (Menschikowski, Hagelgans et al. 2006). A cationic motif (common to sPLA2-IIA and V) allows the enzyme to bind strongly to proteoglycans in the vessel wall (Murakami, Kudo et al. 1989), improving the longevity of the enzymes within the vessel and thus increasing its potential for activity. This is known as the heparan sulphate proteoglycan (HSPG) binding pathway (Figure 1.12).

1.5.2. Secretory phospholipase A2 Group V (sPLA2-V)

The common amino acid sequence for the sPLA2-V protein is shown in Figure 1.11. Secretory PLA2-V is secreted at the subcellular level. It is observed to be most frequently expressed in heart, placenta, macrophages and in lung and other respiratory tissues (http://www.genwaybio.com/product_info.php?products_id=45346).

Secretory PLA2-V is structurally similar to sPLA2-IIA and has been implicated in inflammation as an important component of the cyclooxygenase-2-dependent prostaglandin production pathway on the nuclear envelopes of cells (Balsinde, Shinohara et al. 1999; Kim, Kim et al. 2002). Secretory PLA2-V also plays a role in the arachidonic acid pathway and is considered to be a more potent hydrolyser of lipoproteins than sPLA2-IIA (Lambeau and Gelb 2008). This is probably due to the increased ability of sPLA2-V to catalyse the phospholipids most commonly associated with lipoprotein surface layers; PC and PE. Secretory PLA2-V can hydrolyse both of these with potent effect (Murakami, Kambe et al. 1999; Oestvang, Anthonsen et al. 2003).

Since it has been shown to have a particular efficacy for PC hydrolysis, sPLA2-V is thought to contribute to the hydrolysis of the PC content on HDL surfaces, altering the HDL apolipoprotein A (Apo-A) complex and impeding the reverse cholesterol transport pathway, thus increasing the risk of atherosclerosis (Saiga, Morioka et al. 2001). In fact, sPLA2-V is thought to be more potent towards HDL compared to LDL, due to the increased sphingomyelin found in LDL, which prevents the action of sPLA2s (Murakami and Kudo 2003).

Previously published data (Sawada, Murakami et al. 1999) suggest that when sPLA2-V is co-transfected with cyclooxygenase (COX)-2 (but not COX-1) into human embryonic kidney (HEK) 293 cells it significantly increases the interleukin-1-dependent prostaglandin E2 generation over a 24 hour period of culture. Additionally, rat mastocytoma RBL-2H3 cells overexpressing sPLA2-V showed increased immunoglobulin E-dependent prostaglandin D2 generation and accelerated β -hexosaminidase exocytosis. These results suggest that, like sPLA2-IIA, sPLA2-V acts as a regulator of inflammation-associated cellular responses and though sPLA2-IIA and sPLA2-V appear to work through different pathways (Sawada, Murakami et al. 1999), this is further evidence that there could be some complementary behaviour between them, which may go some way to explaining why strains of mice that are sPLA2-IIA deficient can live normally for an average life-span (Sawada, Murakami et al. 1999).

The sPLA2-V enzymes have the capability of proteoglycan binding and use the HSPG pathway to interact with phospholipids in the same way as sPLA2-IIA (Lambeau and Gelb 2008). However, sPLA2-V has increased interfacial binding capability compared to sPLA2-IIA and thus can also act directly on the phospholipid outer layer of lipoproteins. This is known as the external plasma membrane pathway (Figure 1.12) (Lambeau and Gelb 2008).

1.5.3. Secretory phospholipase A2 Group X (sPLA2-X)

The common amino acid sequence for the sPLA2-X protein is shown below in Figure 1.11. Human sPLA2-X is expressed in various tissues including thymus, spleen, leukocytes, colon and lung tissue. The presence of sPLA2-X in these organs suggests it could be involved in immunity, including inflammation (Gora, Lambeau et al. 2006).

PLA2G2A Translation (144 aa):

PLA2G5 Translation (138 aa):

PLA2G10 Translation (165 aa):

MGPLPVCLPIMLLLLLPSLLLLLLLPGPSGEASRILRVHRRGILELAGTVGCVGPRTPIAYMKYGCFCG
LGHHGQPRDAIDWCCHGHDCCYTRAEEAGCSPKTERYSWQC VNQSVLCGPAENKCQELLCKCDQEIANCL
AOTEYNLKYLFPQFLCEPDSPKCD

Figure 1.11 The protein coding sequences for each of the sPLA2 enzymes: sPLA2-IIA, sPLA2-V and sPLA2-X. In comparison, each protein has a different length amino acid sequence and the lack of significant homology between sequences suggests that each protein has independent actions and none of these three sPLA2 enzymes are redundant in normal physiology (adapted from <http://www.ncbi.nlm.nih.gov/CCDS/> <http://www.ncbi.nlm.nih.gov/CCDS/>, <http://www.ncbi.nlm.nih.gov/CCDS/> and <http://www.ncbi.nlm.nih.gov/CCDS/>).

sPLA2-X has the most potent hydrolytic effects on lipoproteins compared to other sPLA2 enzymes (Hurt-Camejo, Camejo et al. 2000). Secretory PLA2-X is a very effective stimulator of the arachidonic acid pathway (Henderson, Chi et al. 2007) and its activity and high affinity for hydrolysing phospholipids, especially those found in HDL particles to produce arachidonic acid, can be regulated by sphingomyelin and ceramide activity (Oestvang, Anthonsen et al. 2003; Singh, Gesquiere et al. 2007). Increased sphingomyelin was shown to inhibit sPLA2-X activity and to decrease the amount of arachidonic acid produced by sPLA2-X hydrolysis of PC in HDL (Singh and Subbaiah 2007). Conversely, when there is an abundance of ceramide in the environment sPLA2-X activity is increased and there is a significant increase in arachidonic acid as a product of PC hydrolysis by sPLA2-X, especially regarding HDL (Singh and Subbaiah 2007). Secretory PLA2-X has a preferential affinity for

HDL hydrolysis due to its high PC content (Ishimoto, Yamada et al. 2003; Karabina, Brocheriou et al. 2006).

sPLA2-X lacks the cationic motif that allows both group IIA and V to interact with proteoglycans. It is therefore thought that sPLA2-X uses the external plasma membrane pathway (Figure 1.12) to induce the majority of its cellular arachidonic acid release (Bezzine, Bollinger et al. 2002; Pruzanski, Lambeau et al. 2007) and it has been shown to have a strong ability to elicit arachidonic acid production when added exogenously to adherent mammalian cells- something which neither group IIA or V can do (Bezzine, Koduri et al. 2000; Pruzanski, Lambeau et al. 2007).

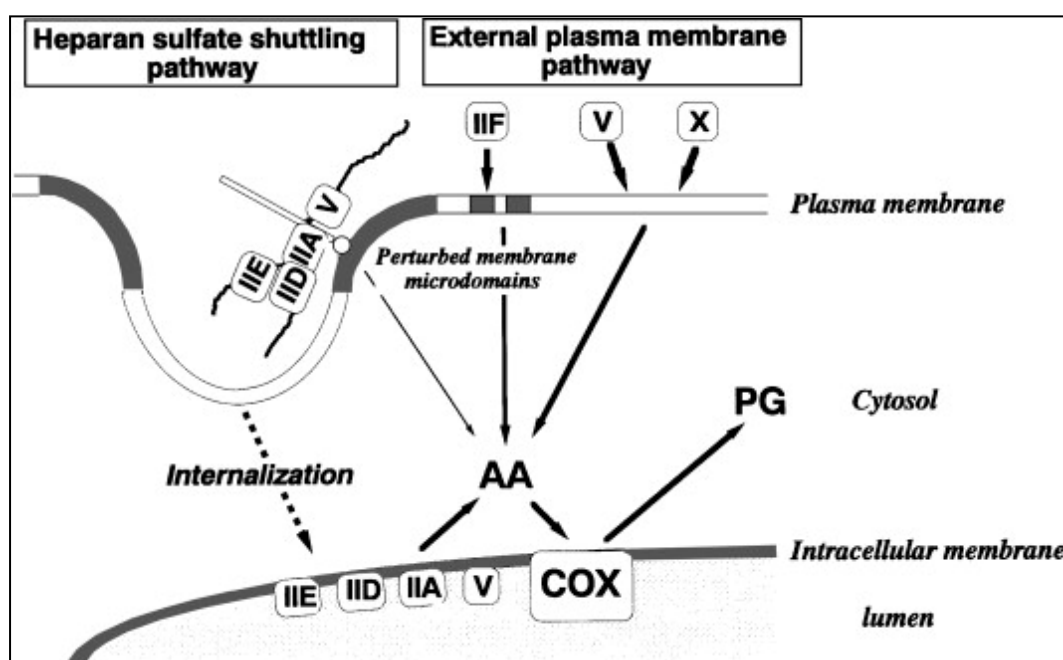


Figure 1.12 The Heparan sulphate pathway (used by sPLA2-IIA and sPLA2-V) and the Plasma Membrane Pathway (Used by sPLA2-V and X) Reprinted with permission from (Kudo and Murakami 2002). The figure shows most of the enzymes in the sPLA2 cluster, sPLA2-IIA, IID, IIE and V are internalised to the intracellular membrane via the Heparan sulphate pathway where they interact with intracellular membrane phospholipids to produce arachidonic acid. sPLA2-IIF interacts with internal membrane microdomain phospholipids directly to produce arachidonic acid. Both sPLA2-V and sPLA2-X use the external plasma membrane pathway to interact directly with the external plasma membrane phospholipids to produce arachidonic acid.

Secretory PLA2-X expression also appears to be activated under inflammatory conditions (Curfs, Ghesquiere et al. 2008). Unlike sPLA2-IIA and V it is also induced by high levels of platelet activating factor (PAF) and plays a role in PAF hydrolysis and regulation (Karabina, Gora et al. 2010). The preferential, precision hydrolysis of PC is what makes

sPLA2-X one of the most potent of the sPLA2 enzymes regarding lipoprotein modification *in vitro*. The selective hydrolysis of PC by sPLA2-X produces a large amount of unsaturated fatty acids and lysoPC (Saiga, Morioka et al. 2001). The enzyme is particularly active when hydrolysing PC on lipoprotein particles and in the outer leaflet of intact cell plasma membranes (Saiga, Morioka et al. 2001). Since HDL is a preferential substrate for sPLA2-X hydrolysis, modifications of HDL due to sPLA2-X hydrolysis means that it is less able to regulate cholesterol efflux, which impedes the reverse cholesterol transport pathway from the arteries to the liver, allowing macrophages to build up and accumulate as foam cells (1.4.5) (Saiga, Morioka et al. 2001).

Since sPLA2-X does not have significant HSPG binding, it is not retained by proteoglycan binding in the extracellular matrix of atherosclerotic plaques, however, this may not be necessary for sPLA2-X to contribute to sPLA2 activity due to its increased potency, dependent on whether there are significant levels of sPLA2-X produced. Despite the increased efficiency of sPLA2-X towards hydrolysis of PC in LDL particles, it has been shown that sPLA2-X-treated LDL particles do not aggregate as sPLA2-IIA modified particles are known to do (Gora, Perret et al. 2009), which is likely due to the additional tryptophan residues found in sPLA2-X compared with sPLA2-IIA and their composition (Gora, Perret et al. 2009). However, these sPLA2-X modified particles are efficiently incorporated into macrophages (Curfs, Ghesquiere et al. 2008), increasing the formation and accumulation of foam cells. Immunohistochemistry has detected sPLA2-X in foam cell lesions the arterial intima and in the smooth muscle cells of the medial layer of the artery wall in *Apoe* deficient mice on an atherosclerotic diet (Hanasaki, Yamada et al. 2002; Murakami and Kudo 2003). This data implicating sPLA2-X in atherosclerosis has been supported by a study designed to detect sPLA2-X in human atherosclerotic lesions from human samples taken from the media and intima (Karabina, Brocheriou et al. 2006) which used double immunohistochemical staining to show that sPLA2-X resides in the atherosclerotic plaque. While there is some evidence to suggest that sPLA2-X may be involved in atherosclerosis, sPLA2-X is not normally expressed in the heart (Masuda, Murakami et al. 2005). In addition to its expression in immune/inflammatory organs and cells, sPLA2-X is expressed most readily in the digestive organs and in the testis. However, it is its expression in blood leukocytes which may explain its presence in the atherosclerotic plaque, since leukocytes are recruited during the inflammatory process of atherogenesis (Kudo and Murakami 2002).

1.5.4. The Genetics of Secretory Phospholipase A2s

The genes that control the production of sPLA2 A2s are among the *PLA2* family of genes. Specifically, sPLA2-IIA, V and X are encoded for by *PLA2G2A*, *PLA2G5* and *PLA2G10*, respectively.

PLA2G2A and *PLA2G5* are tightly linked and transcribed in opposite directions, mapping to chromosome 1p34-p36.1 and only 90Kb apart (Tischfield, Xia et al. 1996). In addition, the genes *PLA2G2D*, *PLA2GE* and *PLA2GF* are all found in a cluster with *PLA2G2A* and *PLA2G5* (Figure 1.13), which could indicate control of all these genes by common promoter elements (Ghesquiere, Gijbels et al. 2005), although this has yet to be proven, since there is no experimental evidence of the precise length of the *PLA2* promoter regions. The percentage of identical amino acids of each gene compared to *PLA2G2A* is shown in Figure 1.13. *PLA2G5* is often co-expressed with *PLA2G2A* in mammalian species in a tissue specific pattern (Tischfield, Xia et al. 1996). There is no observed recombination between the two genes, suggesting they are often inherited together (Tischfield, Xia et al. 1996).

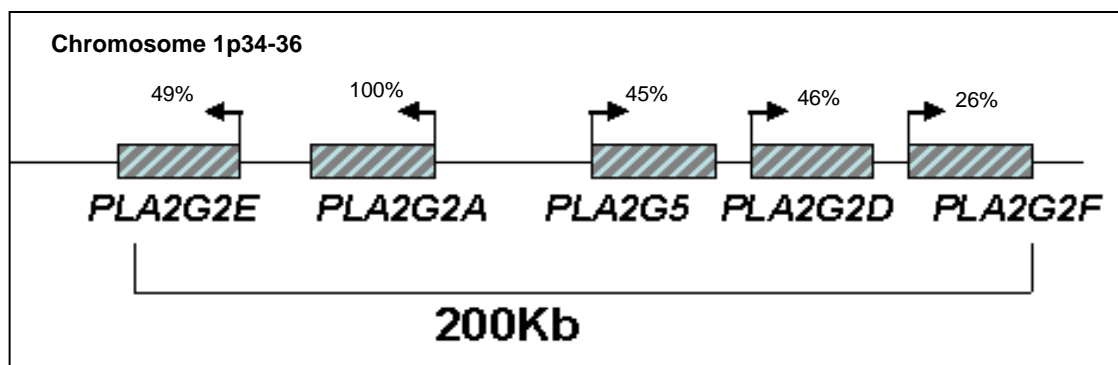


Figure 1.13 Taken from Wootton et al. 2006. Chromosome 1 p34-36, the *PLA2* gene cluster. *PLA2G2A* and *PLA2G5* are transcribed in different directions and are ~90Kb apart. The percentage value above each gene is indicative of the amount of identical amino acids compared to *PLA2G2A*.

The sPLA2-X coding gene, *PLA2G10* but lies on chromosome 16 at 16p13.1-p12 in a gene dense region (Figure 1.14) (Cupillard, Koumanov et al. 1997). In comparison to *PLA2G2A* the *PLA2G10* gene shows only 25% identical amino acid sequence (Ensembl.org).

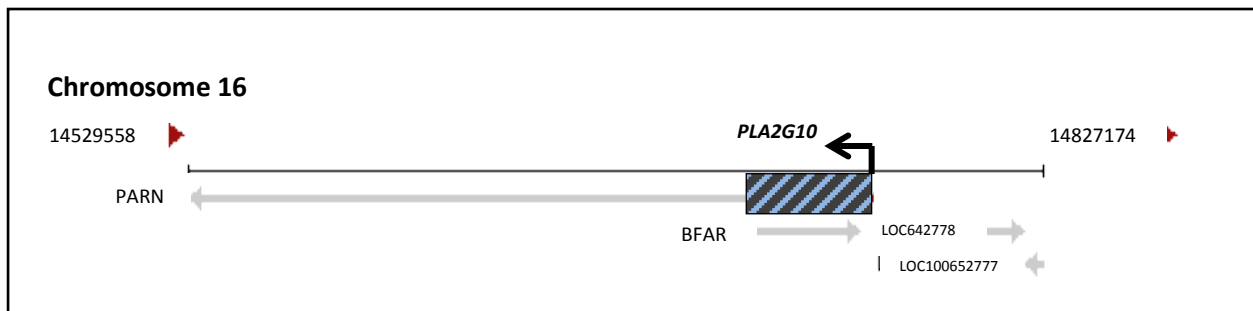


Figure 1.14. *PLA2G10* mapped to its location on chromosome 16, a gene dense region including *PARN*, *BFAR*, *LOC642778* and *LOC100652777*. *PLA2G10* is reverse transcribed as indicated by the black arrow. (<http://www.ncbi.nlm.nih.gov/gene/8399>).

1.5.5. Measuring sPLA2 Mass and sPLA2 Activity

An assay that determines enzyme mass will quantify the amount of enzyme present in a sample matrix (e.g. plasma or synovial fluid), whereas an assay that determines enzyme activity is a measure of the ability of an enzyme to catalyse a specific reaction.

Regarding the enzymes sPLA2-IIA, V and X, development of assays to measure the mass and activity of enzyme has yet to be fully achieved. There is an assay available to measure sPLA2-IIA mass. This is an enzyme-linked immunoabsorbant assay (ELISA), with an sPLA2-IIA specific monoclonal antibody to directly measure the amount of sPLA2-IIA specific antigen in a sample (Koenig, Vossen et al. 2009). This assay is specific for sPLA2-IIA and has been shown to have no cross-over with other known sPLA2 enzymes (Cayman Chemical Co., Ann Arbor, MI, USA), despite similarities in structure with other group II enzymes such as sPLA2-IID, IIE and IIF (Figure 1.9). To date no such assays have been developed to measure either sPLA2-V or sPLA2-X in the same way. It is unclear why no similar assays have been developed for sPLA2-V and X but this may be due to the fact that sPLA2-IIA was the enzyme initially discovered and was of interest to scientists before any more recently discovered isoenzymes including groups V and X. In any case it is surprising that there still appears to be no effort to produce assays for the measurement of sPLA2-V and X that are synonymous with the current sPLA2-IIA assay.

The assay commonly used to measure sPLA2 activity is problematic since the measure of activity obtained includes the individual activity contributions of sPLA2-IIA, sPLA2-V and sPLA2-X. There is currently no assay that can detect each individual input, thus veiling the true value of each of the enzymes' activity. The assay in question is a selective fluorometric assay, where the active enzyme is determined by the production of an increased fluorescent signal (Koenig, Vossen et al. 2009). However, if assays are developed

to measure the individual activity of each enzyme we are still left with the issue of the complexity and precision of the term 'activity.' Measuring the biological activity of any enzyme by an assay using an artificial substrate is only an imperfect model of the substrate formed. Both the sPLA2-IIA mass and the sPLA2 activity assays are used for taking *in vitro* measurements.

1.5.6. Biological Functions of the sPLA2 Family

As previously described, sPLA2 is expressed in response to signals from an inflammatory cascade (**1.5.1-1.5.3**). Some examples of the inflammatory markers known to initiate expression of each enzyme include several cytokines, growth factors and transcription factors. The presence of each enzyme is in a cell and tissue specific manner (Peilot, Rosengren et al. 2000; Menschikowski, Hagelgans et al. 2008). A prolonged inflammatory response drives prolonged *PLA2* gene activation. It is thought that the chief role of sPLA2 enzymes *in vivo* is as part of the hosts defence mechanism against bacterial, viral and parasitic infection (Lambeau and Gelb 2008). Previous studies have shown that sPLA2 is active against both gram positive and gram negative bacteria. Secretory PLA2-IIA was shown to work together with a bactericidal/permeability increasing protein (BPI) to breakdown Gram-negative bacteria such as *Escherichia coli*. The BPI disturbs the surface lipopolysaccharide coat of the *E.coli*, allowing the sPLA2-IIA enzyme to efficiently interact with and break down the cell wall.

Secretory PLA2-IIA was then shown to be most potent against Gram positive bacteria such as *Staphylococcus aureus*, *Streptococcus pyogenes* and *Listeria monocytogenes* (Lambeau and Gelb 2008). No co-enzyme activity is required in this case and the sPLA2-IIA enzyme can act directly on the Gram-positive bacteria cell wall in order to break it down. The physiological properties of sPLA2-IIA are such that the protein is extremely basic and cationic. This charge allows the enzyme to form strong electrostatic interactions with the acidic Gram-positive bacteria surface before binding of the substrate to the active site causes the hydrolysis of the phospholipid membrane (Buckland and Wilton 2000; Lambeau and Gelb 2008).

Other sPLA2 groups exhibit antibacterial activity and a previous study has ranked them in order of potency against Gram-positive activity *in vitro*; group IIA>X>V>XII>IIE>IB/IIF (Koduri, Gronroos et al. 2002). Additionally sPLA2-V was identified as present in the phagosome of macrophages, regulating phagocytosis, suggesting it plays a role in the killing of ingested bacteria (Balestrieri, Hsu et al. 2006).

Antiviral and anti-parasitic roles for sPLA2 groups are less well studied, though it has been shown that sPLA2-X can help to inhibit infection by the human immunodeficiency virus (Fenard, Lambeau et al. 2001; Kim, Chakrabarti et al. 2007) and sPLA2-V and X have been shown to resist adenoviral infection in mammalian cells (Mitsuishi, Masuda et al. 2006). Finally sPLA2-IB can block red blood cell infection by the malaria parasite, *Plasmodium falciparum* (Deregnaucourt and Schrevel 2000; Lambeau and Gelb 2008).

1.5.7. sPLA2s in Disease

Table 1.3 summarises the known and suspected contributions to disease (Pruzanski, Keystone et al. 1988; Henderson, Oslund et al. 2011; Murakami, Taketomi et al. 2011). It appears that each enzyme is expressed differentially depending on the disorder and in some instances all three of the sPLA2 enzymes of interest appear to be associated with a disease (though this is largely based on *in vitro* studies and therefore requires further investigation). This raises the question of whether one or more of these enzymes are redundant or whether differential expression and behaviour of each enzyme contributes to disease. Should the latter be the case then it is also important to explore what the contribution of each enzyme is. As has been discussed previously (**1.1**) CHD is the most significant cause of global mortality and thus it is of great importance to identify causal risk factors associated with this disease. I have therefore chosen to focus my thesis on the investigation of the role of sPLA2 enzymes in CHD, however it is important to also consider the role of these enzymes in other inflammatory disorders, such as rheumatoid arthritis/respiratory inflammation, and in known CHD related diseases such as Type II diabetes (as discussed below). This study will require an understanding of the interactions and contributions of each individual sPLA2 enzyme and this information may also be of use in researching the role of sPLA2 in other disorders.

sPLA2 Enzyme:	Presence of sPLA2 Enzymes in Disease						
	Rheumatoid Arthritis:	CHD†:	ARDs‡:	Asthma:	Sperm Function:	Alzheimers Disease:	Cancer §:
sPLA2-IIA Reference:	✓ (Pruzanski et al., 1988)	✓ (Mallat et al., 2010)	✓ (Nakos et al., 2005)	✓ (Calabrese et al., 2000)	✓ (Takayama et al., 1991)	✓ (Moses et al., 2006)	✓ (Tribler et al., 2007)
sPLA2-V Reference:	✓ (Masuda et al., 2005b)	✓ (Mallat et al., 2010)	✓ (Ohtsuki et al., 2006)				
sPLA2-X Reference:	✓ (Masuda et al., 2005b)	✓ (Mallat et al., 2010)	✓ (Touqui and Alaoui-El- Azher, 2001)	✓ (Henderson et al., 2011)	✓ (Sato et al., 2011)		✓ (Tribler et al., 2007)

† CHD = Coronary heart disease
‡ ARDs= Acute respiratory distress syndrome
§ Cancer includes several forms of cancer including colon cancer and lung cancer

Table 1.3. Highlights the diverse nature of sPLA2 enzyme activity in disease. The differential expression patterns of each enzyme and their distinct catalytic mechanisms suggest independent roles for each enzyme in several diseases. However, in the instance of rheumatoid arthritis, CHD and ARDs all three enzymes have been identified as having a potential association with these diseases. Whether one enzyme is more clinically significant than any other is currently an unanswered question. The evidence for most of these associations relies on *in vitro* data, observational data in small datasets or animal models and more research is required to confirm the existence and potential causality of these associations.

sPLA2-IIA Role in Rheumatoid Arthritis

sPLA2-IIA is considered to play a pro-inflammatory role in rheumatoid arthritis and studies have shown that sPLA2-IIA is highly expressed in synovial cells and chondrocytes in the joints of rheumatoid arthritis sufferers, but not in those of normal subjects (Seilhamer, Pruzanski et al. 1989; Pruzanski and Vadas 1991). This has led to the hypothesis that sPLA2-IIA plays an aggravating role in this autoimmune disease (Murakami, Taketomi et al. 2011). Further support for this theory has come in more recent studies that have utilised the C57BL/6 mouse model, which has a natural frame-shift mutation for *pla2g2a* (Kennedy, Payette et al. 1995). The natural knockout mouse model strain has been used to look at the role of active sPLA2-IIA in inflammation (Ghesquiere, Gijbels et al. 2005) and also in atherosclerosis. These mice are excellent transgenic models for human sPLA2-IIA, as any

recorded changes in atherosclerosis attributed to sPLA2-IIA are due to the human transgene since there is no confounding by endogenous mouse sPLA2-IIA (Webb, Bostrom et al. 2003). Regarding rheumatoid arthritis, these C57BL/6 mice were bred with BALB/c mice (which do not have a mutation for *pla2g2a*) and backcrossed with BALB/c mice. The resulting *pla2g2a*^{-/-} BALB/c offspring were used to show that sPLA2-IIA plays a role in rheumatoid arthritis. This was revealed when arthritis was induced in *pla2g2a*^{+/+} BALB/c and *pla2g2a*^{-/-} BALB/c mice using the K/BxN autoantibody and the levels of inflammatory arthritis were measured in both models. The *pla2g2a*^{-/-} mice showed a significant reduction in arthritic inflammation compared to the *pla2g2a*^{+/+} models (Boilard, Lai et al. 2010). Additionally, *PLA2G2A* transgenic mice have been used with the same K/BxN autoantibody to induce arthritis and increased levels of inflammatory arthritis were seen in these models. It has not yet been confirmed whether sPLA2-IIA contributes to the synthesis of eicosanoids or acts through a different mechanism in inflamed joints (Murakami, Taketomi et al. 2011).

sPLA2-IIA Role in Atherosclerosis and CHD

Observational Data: sPLA2-IIA and Atherosclerosis

Observational studies have been the chief source identifying a relationship between, higher sPLA2 mass and activity and risk of CHD and several studies report this. The first study recruited patients with CAD and sex-matched controls and measured their plasma sPLA2-IIA levels. The study identified the association between elevated PLA2 and increased atherosclerosis/CHD risk. It went on to specify the secretory family of phospholipases as a subgroup associated with this action (Kugiyama, Ota et al. 1999). Studies of the prospective nested case-control study, the European Prospective Investigation of Cancer (EPIC)-Norfolk cohort showed measures of sPLA2-IIA levels have a strong positive correlation with CRP levels and furthermore an independent association with both fatal and non-fatal CHD risk in men (Boekholdt, Keller et al. 2005). The Global Registry of Acute Coronary Events (GRACE) cohort, an observational cohort of acute coronary patients recruited soon after MI, gave results confirming a significant association of sPLA2-IIA levels with risk, as well as showing an association with measures of sPLA2 activity and major recurrent coronary events. The sPLA2 activity measures were a more accurate predictor of death (Mallat, Steg et al. 2005). The same group later used the EPIC study to show a significant association between sPLA2 activity and incidents of CHD (Mallat, Benessiano et al. 2007). A further study examined sPLA2 levels and activity in the Langzeiterfolge der Kardiologischen Anschlussheil-Behandlung (KAROLA) cohort of post myocardial infarction (MI) patients and showed significant associations between sPLA2 levels and activity with fatal and non-fatal MI and stroke (Koenig and Khuseyinova 2009). The association results from all these studies were

equally strong. Table 1.4 gives an overview of the outcome of these and several subsequent studies examining the association between sPLA2 and CHD risk.

Reference	sPLA ₂ Assay	Study Design	Population	No. of Individuals	Average Follow-Up, y	Primary End Point(s)	Risk Estimates (95% CI) ^a
Boekholdt et al and Mallat et al	Type IIA mass or total sPLA ₂ activity	Nested case-control	Apparently healthy men and women	2209 Controls and 1105 cases for mass; 1806 controls and 991 cases for activity	6	Fatal and nonfatal coronary heart disease	OR 1.34 (1.02–1.71) Q4 vs Q1 with mass; OR 1.65 (1.27–2.12) with activity
Kugiyama et al	Type IIA mass	Cohort	Consecutive patients with angiographically proven stable CAD	142 Patients (48 events)	2	Coronary death, MI, and coronary revascularization	OR 3.3 (1.3–9.2) T3 vs T1
Liu et al	Type IIA mass	Cohort	Stable coronary patients undergoing PCI	247 Patients (87 events)	2	Cardiac death, nonfatal MI, repeat angioplasty, or CABG due to recurrent angina	OR 2.1 (1.4–7.0) for sPLA ₂ IIA mass >450 ng/dL
Koenig et al	Type IIA mass or total sPLA ₂ activity	Cohort	Coronary patients 30–70 y old within 3 wk to 3 mo after ACS or CABG	1032 Patients (93 events)	4.1	Cardiovascular death, nonfatal MI, stroke	HR 2.07 (1.17–3.66) T3 vs T1 with mass and 1.65 (0.96–2.84) with activity; HR 1.45 (1.07–1.96) per unit increase with mass and 1.78 (1.03–3.11) per unit increase with activity
O'Donoghue et al	Total sPLA ₂ activity	Cohort	Patients with stable CAD	3738 Patients	4.8	Cardiovascular death, MI, or stroke	HR 1.49 (1.08–2.06) Q4 vs Q1
Kugiyama et al	Type IIA mass	Cohort	Patients with unstable angina	52 Patients (28 events)	2	Coronary death, MI, CABG, PCI, or readmission	OR 5.08 (1.4–18.6) for sPLA ₂ type IIA mass >246 ng/dL
Mallat et al	Type IIA mass or total sPLA ₂ activity	Cohort	Patients with ACS	446 Patients (35 events)	0.54	Death or MI	HR 3.08 (1.37–6.91) T3 vs T1–T2 with activity; HR 1.9 (0.96–3.6) with IIA mass in univariate analysis
Simon et al	Total sPLA ₂ activity	Cohort	Patients with acute MI	1036 Patients (98 events)	0.5	Death, MI, or stroke	HR 2.29 (1.28–4.12) T3 vs T1

Table 1.4. Studies investigating the association of sPLA₂-IIA mass and sPLA₂ activity with cardiovascular event risk (Adapted from (Mallat, Lambeau et al. 2010))

CI indicates confidence interval; OR, odds ratio; Q4, quartile 4; Q1, quartile 1; T3, tertile 3; T1, tertile 1; CABG, coronary artery bypass graft surgery; and HR, hazard ratio.

Fully adjusted risk estimates are depicted for the primary end points.

sPLA₂ mass assay and sPLA₂ activity assay: See **1.5.5**

Animal Models: sPLA2-IIA and Atherosclerosis

The natural knockout mouse model strain for *pla2g2a*; C57BL/6 have also been used as a transgenic model for human *PLA2G2A* when studying sPLA2 association with CHD risk. Mice transgenic for *PLA2G2A* develop atherosclerosis and when macrophages from these mice are transplanted into atherosclerosis-prone knockout mice (either low density receptor *Ldlr*^{-/-} or Apolipoprotein E (*ApoE*)^{-/-} models) atherosclerosis is enhanced (Webb, Bostrom et al. 2003; Boyanovsky, Zack et al. 2009). Specifically, in data from transgenic mice (again strain C57BL/6) over-expressing human sPLA2-IIA it was observed that the enzyme functions in the aorta by altering the conformation of LDL particles, allowing them to aggregate more easily and enhancing their ability to bind to proteoglycans. More specifically it was identified that the apolipoprotein-B100 (ApoB) component of LDL particle was conformationally changed. ApoB has two active sites, designated A and B (among 8 total active sites for proteoglycan binding). Both sites are activated in the presence of sPLA2-IIA and the effects combine to increase proteoglycan binding ability and decreased LDL receptor affinity. However, it has been suggested that it is the conformational change undergone by the site B that causes the greatest enhancement of LDL affinity for proteoglycan binding (Flood, Gustafsson et al. 2004). This has been supported by *in vitro* studies (Eckey, Menschikowski et al. 1997; Jaross, Eckey et al. 2002).

Suggested sPLA2-IIA Pathway to Atherosclerosis

When sPLA2 enzymes hydrolyse the sn-2 ester bond of phospholipids they release a lyso-phospholipid and a non-esterified free fatty acid (NEFA). Release of the NEFA arachidonic acid (AA) is a key step as a precursor in the production of eicosanoids such as leukotrienes, thromboxanes and prostaglandins (Figure 1.8). It therefore promotes these pro-inflammatory lipid mediators which aid the initiation and maintenance of prolonged inflammatory responses in the body, and are implicated in the development of atherosclerosis (Figure 1.16) (Lambeau and Gelb 2008; Leonarduzzi, Gamba et al. 2011).

Studies show that sPLA2-IIA is present in both normal and atherosclerotic arteries, in the circulation and within the vessel wall (Figure 1.16) (Lambeau and Gelb 2008). It is thought to be the most abundant of the three groups of sPLA2 associated with CHD (Mallat, Lambeau et al. 2010) but it is the least potent towards the PC rich phospholipid outer layers of lipoproteins (Lambeau and Gelb 2008; Rosenson and Hurt-Camejo 2012). Though its potency for unmodified LDL is lower compared to sPLA2-V and sPLA2-X, it does have a better affinity for ox-LDL (Lambeau and Gelb 2008; Rosenson and Hurt-Camejo 2012) and its ability to bind strongly to proteoglycans in the extra cellular matrix of smooth muscle cells in the vessel wall, creates an excellent anchor for the longer time it requires to hydrolyse PC

effectively (Murakami, Kudo et al. 1989). As previously discussed, this lack of affinity for PC is probably due to the lack of tryptophan residues compared to sPLA2-V and X, which gives rise to the highly negative charge of sPLA2-IIA enzymes, giving them a very high affinity for binding anionic phospholipid vesicles, but not zwitterionic phospholipid vesicles (Lambeau and Gelb 2008).

sPLA2-IIA expression can be activated by the combined synergistic binding of cAMP and IL-1 β , or by reactive oxygen species (ROS), such as those produced from the NADPH pathway in endothelial cells (Jensen, Sheng et al. 2009) triggering the initiation of *PLA2G2A* transcription (Antonio, Brouillet et al. 2002). In the circulation, hydrolysis of the LDL by sPLA2-IIA modifies the LDL particles generating sd-LDL, which can then transverse the endothelial cell layer of the artery wall into the intima, where they are further modified (Hurt-Camejo, Camejo et al. 2000). This modification increases the propensity of the particles to aggregate and bind to proteoglycans, resulting in aggregation of sd-LDL in the intima. The proinflammatory products released by the sPLA2 hydrolysis of phospholipids stimulate monocytes to enter the intima via the endothelial cell wall where they become macrophages and phagocytose sd-LDL and other sPLA2 modified lipoprotein fragments. This leads to foam cell formation and increased atherosclerotic plaque size (Figure 1.16) (Hurt-Camejo, Camejo et al. 2000).

sPLA2-IIA further modifies LDL in the arterial wall by inducing a conformational change of the ApoB complex in LDL particles, which exposes an extra proteoglycan binding site, promoting the internal aggregation of LDL particles (Tietge, Pratico et al. 2005). This modified LDL is also more susceptible to lipid peroxidation, forming oxidised lipoproteins that bind more tightly to extracellular proteoglycans, increasing atherosclerosis (Tietge, Pratico et al. 2005). The increased levels of foam cells and sPLA2-IIA also appear to contribute to higher oxidative stress within the vessels, increasing inflammatory cytokine expression and therefore further sPLA2-IIA expression in the smooth muscle cells (Tietge, Pratico et al. 2005; Blache, Gautier et al. 2012). The prolonged inflammatory effect also induces smooth muscle cell proliferation, which was cited earlier as a contributing factor to the progression of atherosclerosis (Figure 1.4).

sPLA2-V in Arthritis and Respiratory Disease

An effect of knocking out *pla2g5* in mice (*pla2g5*^{-/-}) leads to an increase K/BxN autoantibody-induced arthritis (Masuda, Murakami et al. 2005). This is probably due to the reduction in the clearance of the pathogenic immune complex by macrophages in inflamed joints in the absence of *pla2g5* (Boilard, Lai et al. 2010). This is validated by a study showing that systemic administration of recombinant sPLA2-V protein promotes the arthritic response

by facilitating the phagocytic uptake of the immune complex by macrophages (Murakami, Taketomi et al. 2011). The pro and anti-inflammatory actions attributed to sPLA2-IIA and sPLA2-V in arthritis are different. This is the likely reason for the ineffectiveness of a combined inhibitor for both sPLA2-IIA and V to combat arthritis, and suggests that drugs specific for each enzyme may be of more use (Murakami, Taketomi et al. 2011).

Regarding respiratory disorders, an association was first noted in 2006 when the first sPLA2 transgenic mice were developed- a *pla2g5* transgenic model. These mice were observed to die soon after birth and the early death was attributed to acute respiratory distress syndrome (ARDS), which is brought on by the abnormal hydrolysis of pulmonary surfactant phospholipids. Pulmonary surfactant is produced by alveolar type II epithelial cells and is essential for maintaining alveolar stability. Destruction of surfactant lipids by high levels of sPLA2-V can cause loss of alveolar stability and impairment of gas exchange (Ohtsuki, Taketomi et al. 2006). Secretory PLA2-V has been shown to be expressed in the bronchial endothelium and alveolar macrophages and over expression of this enzyme is associated with patients with severe pneumonia (Masuda, Murakami et al. 2005; Munoz, Meliton et al. 2007). Further studies have utilised both *PLA2G5* transgenic mouse models and *pla2g5*^{-/-} mouse models and the resulting data indicates that sPLA2-V is involved in both acute and chronic respiratory disorders, by increasing airway inflammation and increased hydrolysis of pulmonary surfactant lipids (Munoz, Meliton et al. 2007; Munoz, Meliton et al. 2009).

sPLA2-V in Atherosclerosis and CHD

Like sPLA2-IIA, sPLA2-V is present in both the normal artery and in the atherosclerotic vessel in an *in vitro* study (Masuda, Murakami et al. 2005). Some studies have failed to find evidence of significant sPLA2-V activity or levels in atherosclerotic lesions, suggesting that this protein has a less important role to play in CAD than sPLA2-IIA (Ishimoto, Yamada et al. 2003). These results are disputed in studies using immunohistochemistry that showed that sPLA2-V is present in the smooth muscle cell associated atherosclerotic lesions of mice and humans as well as in the areas around foam cells (Rosengren, Peilot et al. 2006; Bostrom, Boyanovsky et al. 2007). This suggests that the contribution of sPLA2-V to atherosclerosis may be localised to the intima of the vessel wall, probably altering lipoprotein/proteoglycan interactions at the surface of macrophages and increasing the deposition of collagen (Bostrom, Boyanovsky et al. 2007).

Unlike sPLA2-IIA, both sPLA2-V and X have been shown to be at least 20-fold more efficient in hydrolysing PC, including that seen in LDL and HDL phospholipid outer layers,

due to their high affinity for both anionic and zwitterionic phospholipid vesicles (Lambeau and Gelb 2008).

As previously mentioned, sPLA2-V has a strong effect on HDL. Secretory PLA2-V has been shown to preferably bind to the outer layer of PC in HDL as a substrate, altering the HDL apolipoprotein A (Apo-A) complex and impeding the reverse cholesterol transport pathway (1.4.5) to increase the risk of atherosclerosis (Ishimoto, Yamada et al. 2003).

A study using C57BL/6 mice to create *pla2g5* knockouts from embryonic stem cells that had been transfected with a linearized gene targeting construct, monitored zymosan-induced inflammatory products made in the cells of the wild type (shown to be at high levels) versus the cells of C57BL/6 *pla2gv*^{-/-} mice (shown to be at low levels) and suggested that sPLA2-V has a role in eicosanoid generation regulation in response to an acute innate response of the immune system. This eicosanoid regulation is thought to be in the induced arachidonic acid pathway. Further findings confirmed sPLA2-V promotes the atherogenic process by both liberating potent pro-inflammatory lipid mediators and by modifying LDL to pro-atherogenic LDL (Satake, Diaz et al. 2004).

A likely conclusion is therefore that sPLA2-V does play a role in CHD, but that it acts in a different way to sPLA2-IIA and seems to adopt elements of activity similar to both sPLA2-IIA and X, which have very different actions and activity (Webb, Bostrom et al. 2003; Rosengren, Peilot et al. 2006). However it is important to interpret these results with caution since they are mostly based on *in vitro* assays.

sPLA2-X in Respiratory Disease

Like sPLA2-V, sPLA2-X has been shown to play a pathological role in respiratory diseases such as ARDs and asthma. It is thought to contribute to airway injury through the hydrolysis of pulmonary surfactant lipids, such as PC, in the same way as sPLA2-V (Curfs, Ghesquiere et al. 2008). Studies have also acknowledged the expression of sPLA2-X in the airway epithelial cells of both mice and humans (Masuda, Murakami et al. 2005; Hallstrand, Chi et al. 2007). *PLA2G10* transgenic mouse models and knockout (*pla2g10*^{-/-}) mouse models have been used since 2007 (Henderson, Chi et al. 2007; Curfs, Ghesquiere et al. 2008) to confirm this pathologic effect of sPLA2-X. The use of sPLA2 inhibitors has been shown to improve the condition of ARDS models (Furue, Mikawa et al. 2001; Touqui and Alaoui-El-Azher 2001).

sPLA2-X in Atherosclerosis and CHD

Unlike sPLA2-IIA and sPLA2-V, sPLA2-X is not seen in the healthy human artery, but it has been shown to be present in atherosclerotic plaques by immunohistochemistry (Kimura-Matsumoto, Ishikawa et al. 2008). As previously highlighted, it is the most potent of the three enzymes for hydrolysing PC and like sPLA2-V it has a high affinity for both anionic and zwitterionic phospholipid vesicles (Lambeau and Gelb 2008).

Treating cells with sPLA2-X induces an increase in the negative charge of LDL without the excessive modification linked with oxidation (Hanasaki, Yamada et al. 2002). This process of lowering the charge of LDL particles suggests that sPLA2-X treatment is a very efficient way to complete the prerequisite step required for LDL to promote the formation of foam cells (usually this occurs due to oxidation) and this promotes the pathogenesis of atherosclerotic lesions (Ishimoto, Yamada et al. 2003; Boyanovsky, van der Westhuyzen et al. 2005).

Watanabe heritable hyperlipidemic (WHHL) rabbits were used to show that sPLA2-X (as well as sPLA2-V) unlike some other sPLA2 groups, including IIA, can induce lipolytic modification of HDL (Ishimoto, Yamada et al. 2003). As previously discussed, this reduces the ability of HDL to initiate cellular cholesterol efflux from macrophages containing high lipid levels (**1.4.5**). The study used immunohistochemical analysis to look at WHHL rabbit arterial intima, and observed that there is a significant expression of sPLA2-X in foam cell lesions there (Ishimoto, Yamada et al. 2003). The *PLA2G10* transgenic mouse model and the knockout *pla2g10* mouse model used previously to study the role of sPLA2-X in respiratory disorders (Henderson, Chi et al. 2007) have not currently been used to look at CHD risk.

1.5.8. An Overview of sPLA2 Actions

Secretory phospholipase A2s have been associated with several biological processes and human diseases and the most significant of these have been discussed above. Figure 1.15 has been adapted from a recent review and is a useful schematic tool for presenting an overview of these processes (Murakami, Taketomi et al. 2011). Figure 1.15 is also a good illustration of the many unanswered questions that remain about *in vivo* sPLA2 actions, several of which are addressed in this thesis.

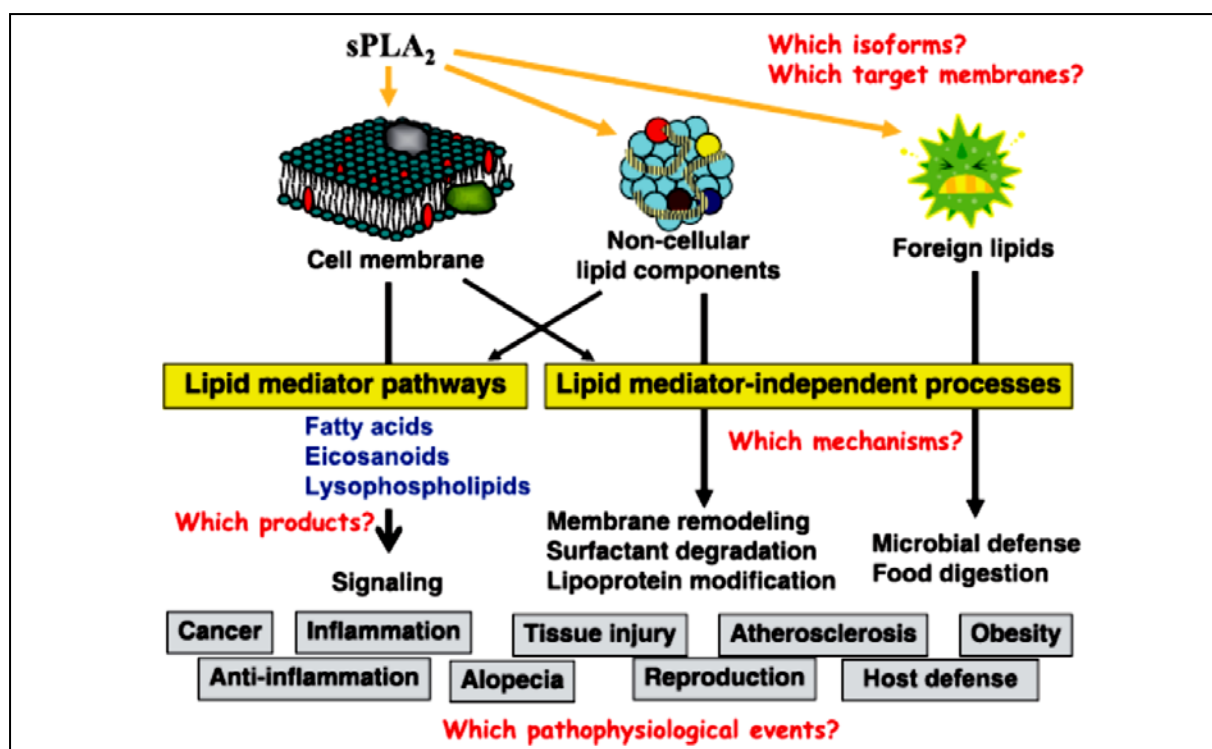


Figure 1.15. A schematic representation of sPLA₂ pathways. This figure represents the known and suspected biological pathways, processes and diseases that sPLA₂ isoforms have been linked with. The figure also indicates how we have not yet comprehensively identified which of the sPLA₂ isoforms contributes to which processes, by what mechanisms and with what outcomes. Reprinted with permission from (Murakami, Taketomi et al. 2011).

Summary of the sPLA₂ Association with CHD Based on Current Data

The identification of elevated sPLA₂-IIA in the serum of rheumatoid arthritis patients (Seilhamer, Pruzanski et al. 1989) raised the question of the role of this enzyme in the inflammatory pathway. Atherosclerosis, which is often the initial stage of CHD, is considered to be an inflammatory disorder, leading researchers to consider the role of the supposed inflammatory biomarker, sPLA₂-IIA, in the onset of CHD. Later, sPLA₂-V and sPLA₂-X were identified as additional potential causal risk factors of CHD. The literature to date has established that all three enzymes exist in the human atherosclerotic plaque (Karabina, Brocheriou et al. 2006; Rosenson and Gelb 2009). However, only sPLA₂-IIA and sPLA₂-V were shown to be present in the normal human artery (Karabina, Brocheriou et al. 2006). It has also been confirmed in observational studies that increased sPLA₂-IIA levels are significantly associated with primary CHD, as has sPLA₂-activity (Koenig and Khuseyinova 2009). Additionally, serum measurements of both sPLA₂-IIA levels and sPLA₂ activity have been shown in observational studies to be associated with secondary events (Lind, Simon et al. 2012). There are however, some remaining questions, which are to be explored in this

thesis. For example, the observational, *in vitro* and *in vivo* data confirm a significant association between sPLA2 enzymes and CHD, probably through the previously established sPLA2/atherosclerosis pathway (summarised below in Figure 1.16). However, we do not know whether this association is causal or whether sPLA2 is merely a marker of atherosclerosis, induced by the onset of disease. In addition there are no known assays to measure sPLA2-V and sPLA2-X mass individually in the serum as we have seen for sPLA2-IIA and there is no assay to measure individual serum activity measures of any of the three enzymes since the current assay is a composite measure of all three enzyme contributions (Mallat, Benessiano et al. 2007). This is problematic since we cannot be sure of which enzyme is the most effective. The group IIA enzyme is thought to be the most abundant (Mallat, Lambeau et al. 2010), but the group V and X enzymes are more potent so it is impossible to know which is the most effective without individual activity measures.

Figure 1.16 depicts the pathways through which previous research has suggested that sPLA2 enzymes contribute to atherosclerosis. It is not known whether these actions are causal for atherosclerosis onset or merely a disease/risk-factor induced marker.

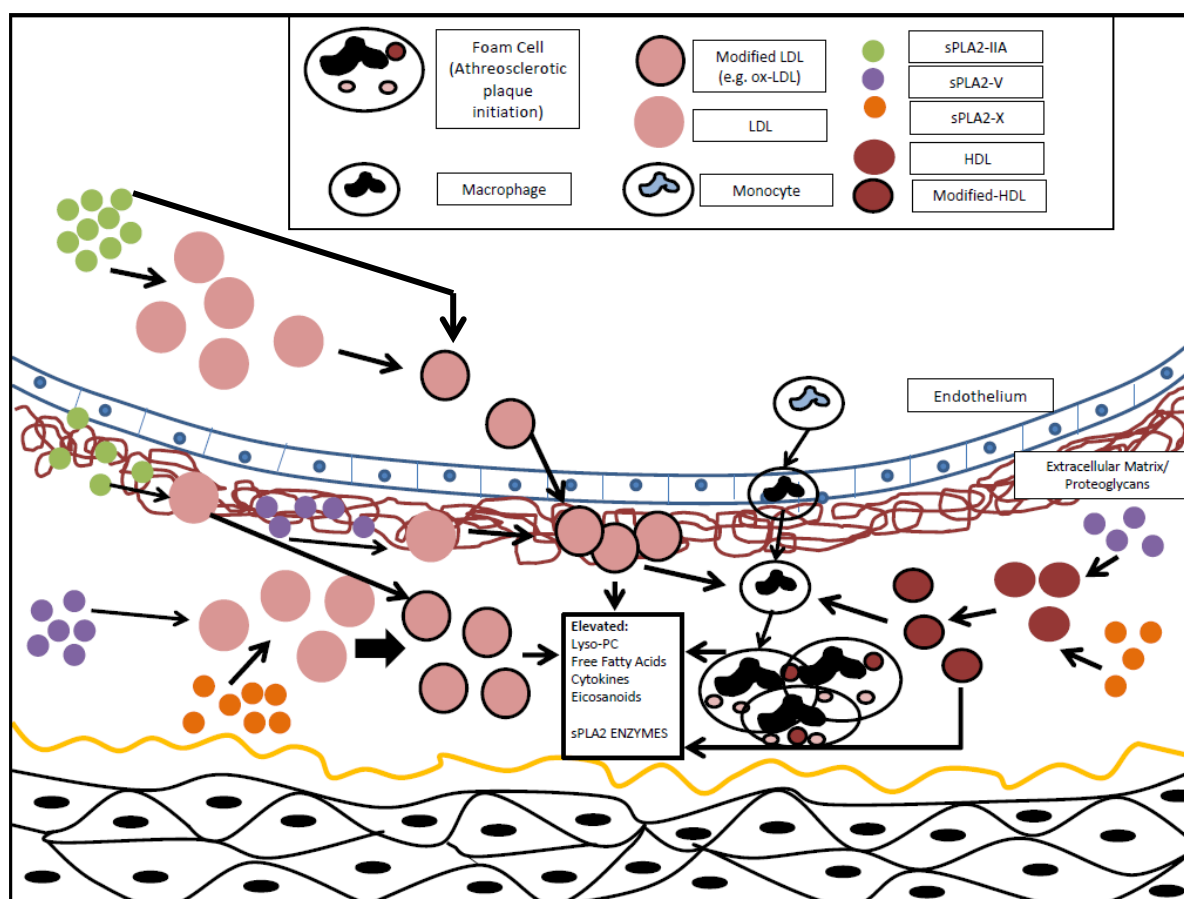


Figure 1.16. Secretory phospholipase interaction with lipoproteins is thought to increase the risk of atherosclerosis. sPLA2 enzymes are induced as part of the inflammatory pathway, which is known to initiate atherosclerosis. However, we do not know whether this pathway is part of the initial causal inflammatory response that induces atherosclerosis, or whether it is a secondary inflammatory response induced by plaque formation or an increase in a known causal biomarker. Secretory phospholipase A2 group IIA (green) is expressed in both the circulation and in the intima. It is thought to be the most abundant of the sPLA2 enzymes, but the least potent towards PC rich lipoproteins. It acts with weak effect on LDL-cholesterol in the circulation and in the extracellular matrix where it has the ability to bind to proteoglycans and modifies LDL to sd-LDL. It acts with slightly more potent effect on modified LDL (for example ox-LDL). Group V sPLA2 (purple) acts on LDL/LDL modified particles in the intima via interactions with proteoglycans in the extracellular matrix and directly on the surface phospholipids of LDL in the intima. Group V sPLA2 also acts on HDL/HDL-modified particles in the intima. As with sPLA2-V, sPLA2-X (orange) acts directly on the surface phospholipids on both LDL/LDL-modified and HDL/HDL-modified particles in the intima, however it cannot interact with proteoglycans to hydrolyse LDL surface phospholipids in the extracellular matrix. sPLA2-V and X are the most potent hydrolysers of lipoproteins such as LDL and HDL, which are high in PC. Modified lipoproteins remain in the intima and have an increased ability to bind to proteoglycans in the extracellular matrix and to aggregate together. Further products of sPLA2 hydrolysis include elevated lyso-phospholipid, free fatty acids, cytokines and increased sPLA2 enzymes. These products induce monocytes to move into the intima and become macrophages. The macrophages begin to phagocytose the modified lipoprotein aggregates and eventually become foam cells, which form the beginnings of atherosclerotic plaques. The increased inflammatory products act to increase sPLA2 enzyme production.

The pathways via which sPLA2 enzymes are thought to contribute to atherosclerosis do not confirm or refute the causality of sPLA2 in the onset of atherosclerosis. Since the enzymes are induced by various inflammatory markers there is currently no definitive evidence that sPLA2 enzymes are part of the inflammatory cascade that initiate atherogenesis or as part of a secondary cascade induced by the onset of the disease. The causality of sPLA2 enzymes in the process of atherogenesis is an important question to address to understand the value of sPLA2s as a candidate drug target for CHD.

1.6. Varespladib: An sPLA2 Inhibitor Under Trial

Inhibitor drug compounds are often designed based on observational and early animal studies where data has indicated that blocking the pathway of a biomarker of interest will be beneficial regarding a disease outcome. An example of a class of drugs that were trialled in this way, with generally excellent outcomes in the reduction of disease versus potential side-effects, is statins. One of the early statin compounds is simvastatin, an inhibitor of HMG-CoA reductase that leads to a decrease of LDL-cholesterol, which was tested to evaluate its suitability for combating death by CHD in a large scale randomised control trial (RCT); The Scandinavian Simvastatin Survival Study (4S) (Kjekshus and Pedersen 1995). This trial showed that lowering LDL-cholesterol by inhibiting the HMG-CoA reductase pathway leads to a reduction in death by CHD. The trial was based on 4,444 patients with angina pectoris or previous myocardial infarction and serum cholesterol levels of 213-310mg/dl while being treated with a lipid-lowering diet. Participants were randomly and blindly allocated to treatment with simvastatin or placebo with a follow-up over 5.4 years. The trial showed that simvastatin was associated with a risk reduction of 34% ($P=0.0001$) with major coronary events. The data showed how important drug inhibitors can be in the prevention and treatment of disease (Kjekshus and Pedersen 1995). This trial clearly confirmed that the relationship between LDL-cholesterol and CHD was a causal one.

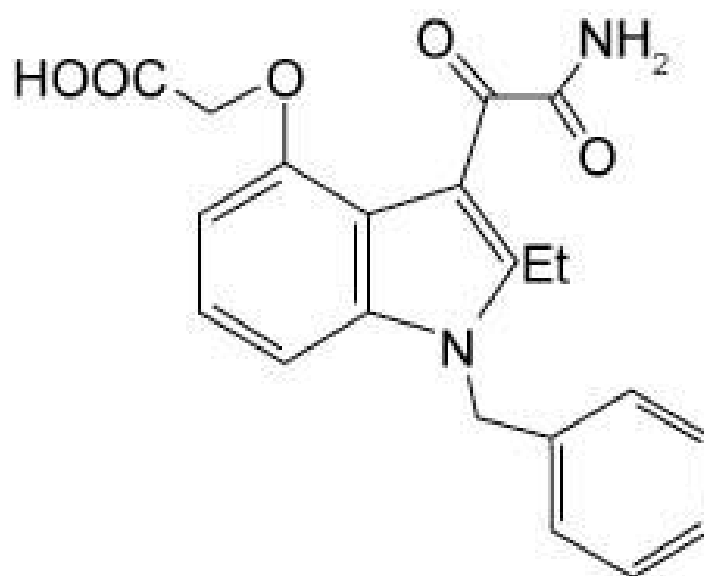
However, it has also been shown that inhibition of biomarkers that look potentially important in the causality of disease can have either unforeseen side-effects or can even be proved not to be causal. For example; Lp-PLA2 was identified as a potential therapeutic target for reducing the risk of CHD since observational studies showed a significant association between Lp-PLA2 activity and increased CHD risk (Kleber, Wolfert et al. 2011). Darespladib is an Lp-PLA2 inhibitory compound designed to decrease the risk of CHD. Animal studies and Phase I clinical trials proved the inhibitory effects of the drug on Lp-PLA2 activity (Garcia-Garcia and Serruys 2009). However, no significant reduction in CHD risk was seen. Additionally, a study investigating genetic variants significantly associated with Lp-PLA2 activity examined the SNP rs1051931 (A379V), which was shown to be the SNP

with the strongest association to Lp-PLA2 activity. The VV genotypes showed 7.2% higher activity compared to AA genotypes. This SNP was not shown to be independently associated with CHD risk, suggesting that Lp-PLA2 does not play a causal role in CHD and therefore inhibition of this enzyme is unlikely to decrease CHD risk (Casas, Ninio et al. 2010). Varespladib is a drug currently under development and comes from the same class of drugs as darespladib, but rather than inhibiting Lp-PLA2, it is a specific inhibitor of sPLA2, the novel biomarker explored during this thesis. Over 40 structural classes of sPLA2 inhibitors exist and all but a few lack potency against sPLA2 action with 50% or more inhibitory concentration (IC_{50}) value (Lambeau and Gelb 2008). Among the known inhibitors being developed is a class of sPLA2 inhibitors known as indoles, including varespladib the first in class to inhibit secretory phospholipases (Lambeau and Gelb 2008). At the beginning of my research, secondary trials for this drug were in their infancy, increasing the interest and validity of my research.

1.6.1. Varespladib Structure and Production

Varespladib (A-002) is an orally delivered drug that was being developed by Anthera Pharmaceuticals (<http://www.anthera.com/>) in a Phase III trial as specific inhibitor of sPLA2-IIA, V and X (Nicholls, Cavender et al. 2011), hypothesising that inhibition of these enzymes will significantly reduce the risk of CHD.

The molecule was first identified by a high throughput screen to obtain lead compounds of interest for sPLA2 inhibition based on structure (Schevitz, Bach et al. 1995; Lambeau and Gelb 2008). Varespladib is a broad sPLA2-inhibitor with IC_{50} values in the nanomolar range. Studies show that these indole compounds contain amide carbonyl and carboxylate groups that act as direct ligands to the active calcium site (including the unique catalytic dyad comprising aspartic acid (Asp) and histidine (His)) of sPLA2 (Mihelich and Schevitz 1999; Smart, Pan et al. 2004; Lambeau and Gelb 2008). The chemical structure for varespladib is shown below in Figure 1.17.



Varespladib

Figure 1.17. The chemical structure of varespladib (Reprinted with permission from (Lambeau and Gelb 2008)). Varespladib is a specific sPLA2 inhibitor drug, which was trialled as a drug for the prevention of secondary cardiovascular events.

Phase I and II trials provided some evidence that the drug may be of use for secondary prevention of major adverse coronary events such as myocardial infarction (Rosenson, Hislop et al. 2010; Rosenson, Elliott et al. 2011) and it was being trialled on a large ACS cohort in the Phase III trial, Vascular Inflammation Suppression to Treat Acute Coronary Syndrome (VISTA-16) (Nicholls, Cavender et al. 2011).

Phase I trials studied the effects of varespladib methyl on *Apoe* knockout strain of atherosclerosis prone mice. The *Apoe* ^{-/-} mice used in these studies had a C57BL/6J background, where a frame-shift mutation causes the loss of sPLA2-IIA production. Varespladib was shown to reduce atherosclerosis in these mice by 48-52% at a dose of 30 and 90mg/kg twice a day for 16 weeks, suggesting that varespladib was a potentially useful anti-atherosclerosis treatment and that may inhibit sPLA2-V and sPLA2-X (since sPLA2-IIA is not present), which implies that these isoenzymes may have additional significant effects towards atherosclerosis (Rosenson, Fraser et al. 2010). Total plasma cholesterol was also reduced by around 22% to 24%. Experiments in *Apoe* accelerated atherosclerosis and aneurysm models, administering a dose of 30mg/Kg twice a day, showed a similar reduction

of atherosclerosis and total elimination of any aneurysm formation (Rosenson, Fraser et al. 2010). An additional study investigated the effects of a combination therapy of varespladib methyl and pravastatin in the *Apoe*^{-/-} model (Shaposhnik, Wang et al. 2009). The results were interpreted to show a synergistic action between pravastatin and varespladib to reduce atherosclerosis, with the additional benefit that varespladib appeared to alter the structure of the lesion to a less mature version with a larger and more stable fibrous cap, thus reducing the likelihood of plaque rupture. In this instance the plasma LDL-cholesterol levels were unaltered by varespladib, but the results did show a 40% increase in HDL and an 80% increase in plasma paraoxonase (PON) levels, both of which are considered cardioprotective (Gordon, Probstfield et al. 1989; Assmann, Schulte et al. 1996; Shih, Xia et al. 2000; Hase, Tanaka et al. 2002).

High-fat, high-cholesterol diets fed to guinea pig models induced atherosclerosis (Leite, Vaishnav et al. 2009). Varespladib was introduced at 150mg/Kg per day and significantly reduced measures of IL-2, IL-10 and IL-12 as well as reducing cholesterol accumulation in aortic tissues by 26.5%. No significant effects were observed on plasma cholesterol levels. Atherosclerosis measured by severity score was decreased by 24% and when measured by digital imaging was reduced by 37%, though neither result was statistically significant. It was also acknowledged that varespladib decreased hepatic lipodosis (fatty liver) by around a third (Leite, Vaishnav et al. 2009).

Phase II trial in humans were named Phospholipase Levels and Serological Markers of Atherosclerosis (PLASMA) I and II, Fewer Recurrent Acute coronary events with Near-term Cardiovascular Inflammatory Suppression (FRANCIS) and the sPLA2 Inhibition to Decrease Enzyme Release after Pre-Cutaneous Intervention (SPIDER-PCI) trial (Rosenson, Fraser et al. 2010; Rosenson, Hislop et al. 2010). PLASMA I was a randomised, double-blind, placebo-controlled parallel arm dose-ranging study (Rosenson, Fraser et al. 2010). Doses were administered twice a day at 50, 100, 250 and 500mg per day in 393 patients with stable CHD for 8 weeks (Rosenson, Hislop et al. 2009). There was a dose-dependent, significant reduction of sPLA2-IIA levels with varespladib (69.2-95.8%, $P=0.0001$). Comparison measures of sPLA2 activity could not be taken due to the enzyme mass in treated patients falling below the threshold necessary for sPLA2 activity measures. Varespladib treatment also appeared to modify LDL particles from small to large, significantly reducing the total LDL-cholesterol compared to placebo as well as decreasing ox-LDL and high-sensitivity CRP (Rosenson, Hislop et al. 2009).

PLASMA-II was a randomised and placebo controlled trial where doses of 250mg and 500mg were administered to 135 stable CHD patients once a day for 8 weeks

(Rosenson, Elliott et al. 2011). The outcomes of this trial were very similar to those of PLASMA-I. A significant reduction in sPLA2-IIA levels was seen in a dose dependent manner. Reductions of 73.4% with 250mg and 84.4% with 500mg were observed. Additionally, LDL-cholesterol was reduced by 12-18% with a shift from smaller to larger particles appearing to drive this reduction and decreased ox-LDL and high-sensitivity CRP levels were also observed (Rosenson, Elliott et al. 2011).

Based on evidence from these trials and previous *in vivo* and observational data, it appeared that sPLA2 acted through the inflammatory pathway to affect secondary coronary event incidence, and was unlikely to have an effect on primary cardiac event incidence. For this reason, varespladib was considered to be a drug that could be effective as a treatment for acute coronary syndrome (ACS) patients, who suffer from increased inflammation post-event, as well as an increased risk of a second cardiac event (Blake and Ridker 2003). Varespladib was considered as an intervention to reduce inflammation and thus also significantly reduce the likelihood of a secondary major adverse coronary event (MACE), such as myocardial infarction, unstable angina or stroke.

The FRANCIS trial was set up to observe the effect of varespladib on sPLA2-IIA levels in patients with acute coronary syndrome (ACS) as sPLA2-IIA was thought to be predictive for future events in this group (Rosenson, Hislop et al. 2010). There were 624 patients enrolled in the trial, all of whom were already being treated with 80mg/day of atorvastatin. The patients were randomised to either a 500mg/day dose of varespladib or a placebo for a minimum of 6 months. Varespladib was shown to significantly reduce sPLA2-IIA levels along with LDL-cholesterol and inflammatory biomarkers. A positive trend was seen for the reduction of unstable angina and myocardial infarction (Rosenson, Hislop et al. 2010).

The SPIDER-PCI trial included 144 patients randomised to either 500mg per day of orally administered varespladib or a placebo 3-5 days before an elective percutaneous coronary intervention (PCI) and then for 5 days following the operation. Results showed that while varespladib significantly reduced sPLA2-IIA levels, no other effects were seen for relevant biomarkers (Dzavik, Lavi et al. 2010).

Despite some positive evidence in Phase I and II trials there were still many questions to be answered as to how effective this drug is at lowering secondary CHD risk, especially in comparison to treatments already in place. At the start of my studies for this thesis, the phase III VISTA-16 trial was in progress. This was a time consuming and expensive trial designed based on phase I and II results. I hope to show how the use of genetics earlier on in the drug development process can add valuable data to the suitability

of an enzyme as a drug target- in this case examining the causality of sPLA2 enzymes for CHD. This process may save time and resources for drug companies.

1.7. Genetic Variations and Disease

The Human Genome Project was completed in 2003 (International Human Genome Sequencing 2004) and findings showed approximately 3.3 billion DNA base pairs along with around 23,000 protein-coding genes. This was far less than previously expected. It is estimated that around 1.5% of the genome is protein coding and the remainder is made up of non-coding ribonucleic acid (RNA) genes, introns, regulatory sequences and 'junk' DNA (International Human Genome Sequencing 2004). There has been an abundance of data available regarding the variation of the genome (Sachidanandam, Weissman et al. 2001) and a great many human DNA sequence variations have been identified due to advances in sequencing technology, which have increased both the speed and efficiency of sequencing. Every genome has around 4 million DNA sequence variants and several different types of variant exist. Some of the most common and important DNA variations are known as single nucleotide polymorphisms (SNPs). There are thought to be approximately 3 million SNPs within the human genome. There are also structural variations such as copy number variants (CNVs) that cover a wide number of nucleotides.

More recently the Encyclopedia of DNA Elements (ENCODE) project has examined the non-protein coding regions of the genome (Consortium, Bernstein et al. 2012). This project aimed to identify the functions of these regions, which make up the vast majority of the human genome. Systematic mapping of the various regions of transcription, transcription factor association, chromatin structure and histone modification was undertaken. The results suggested that over 80% of the genome can now be assigned a biochemical function. Many of these regions have subsequently been identified as candidate regulatory elements and are physically associated with each other and expressed genes and this gives researchers fresh insight into the mechanisms of gene regulation. The statistical correspondence of these newly identified elements with sequence variants was often significant and will aid the interpretation of these variations. The advances made by this project provide new insight into the organization and regulation of our genes and genome and is an excellent resource for future studies (Consortium, Bernstein et al. 2012).

Genetic variations have been shown to have a strong biological and clinical impact. Complex common disorders are often due to many intricate interactions between a large number of genetic variants and environmental factors. Each factor contributes a small effect which can often be clinically insignificant. However, some SNPs contribute a much larger

effect size and can even have clinical significance which could lead to an improvement in treatment, early diagnosis or risk stratification (Marian 2010).

For this thesis it was important to identify SNPs that played a functional role in sPLA2-IIA to use as viable genetic instruments in a large scale meta-analysis. By attempting to identify functional SNPs that showed differential effects on sPLA2-IIA mass and/or activity and the pathways they worked through, it would be possible to confidently predict that any associations with either CHD or CHD biomarkers seen with these variants would be truly due to changes in sPLA2-IIA levels or sPLA2 activity. The use of tagging-SNP (tSNP) analysis of *PLA2G2A* demonstrated a strong association of serum levels of sPLA2-IIA and *PLA2G2A* variants (Wootton, Drenos et al. 2006). Results also showed levels of sPLA2-IIA were negatively correlated with total antioxidant status and low HDL cholesterol in men (independent of CRP) and in women levels of sPLA2-IIA mass were correlated with CAD status. These results highlighted for the first time the relationship between genetics, elevated sPLA2-IIA levels and CAD (Wootton, Drenos et al. 2006).

1.7.1. Genome-Wide Association Scans

Rather than focusing on a small candidate gene area to identify possible clinically relevant variations, a genome wide association scan (GWAS) study will scan the whole genome to detect all variations included in the array. It is a hypothesis free, unbiased approach allowing a comprehensive, genome wide survey. The limitations of such studies include the need for very large sample sizes, to avoid multiple testing errors (Suh and Vijg 2005) and should a disease associated variant occur at a low frequency it will often be undetected by GWAS. It is also relatively expensive to carry out.

1.7.2. Indirect/Direct Candidate Gene Approach

This is an alternative approach to GWAS that utilizes the indirect/direct candidate gene pathway, based on prior knowledge of the disease phenotype and genetic studies in model organisms or genetic location studies (Tabor, Risch et al. 2002). Pre-defined tagging-SNPs (tSNPs) that exist across a gene or region of interest are utilised in indirect candidate association studies. Tagging-SNPs are variants that have been shown to occur in high linkage disequilibrium (LD). This means that the likelihood of a combination of variations or genetic markers occurring more or less frequently in a population compared to what would be expected by chance, i.e. there is little recombination between these alleles. Tagging-SNPs can therefore be used to analyse large areas of a gene, since one tSNP is able to represent all SNPs in high LD with it across the gene. High LD is considered to be calculated at >90%. LD can be represented by either a D' or an R^2 value. The D' value is a

measurement of LD between two variants that does not take into account the frequencies of each variant. Therefore two SNPs may be in high D' but one allele occurs at a much higher frequency in a population than the other. A more precise LD measurement is R^2 , which does take into account allele frequency and therefore means that a high value not only suggests that two variants are often inherited together in a population, but additionally that they occur at a similar frequency in that population. Studies using tSNPs are used to investigate the genetic aetiology of complex diseases such as CHD. This approach was used to identify the tagging SNPs associated with the *PLA2* genes.

A previous study has identified tagging-SNPs for *PLA2G2A* (Wootton, Drenos et al. 2006). Any significant associations between biomarkers and a tSNP can reasonably be assumed to occur with all tagged SNPs. Therefore tSNPs allow us to potentially interpret the effects of SNPs capturing the majority of the variance of a gene of interest in a cost and time-effective manner. However, once a significant association between a tSNP and a biomarker has been identified the question is raised about which SNP is the true functional variant, the tSNP or a tagged variant(s)? Knowledge of which SNP is functional can be important when attempting to assess the biological mechanism through which the variant acts, potentially helping to identify new drug targets. It also provides useful genetic tools for large scale observational studies. For the *PLA2G2A* gene, the six identified tSNPs captured 92% of the variation of *PLA2G2A* (Wootton, Drenos et al. 2006). These tSNPs were genotyped in a cohort of patients with type 2 diabetes (the UDACS cohort- See **2.5**) and their association with sPLA2-IIA mass and lipid traits was explored. Baseline measurements for sPLA2-IIA levels were taken in UDACS samples with and without CAD; 3.08ng/ml in samples without CAD (n=383) and 3.45ng/ml in patients with CAD (n=136). Additional baseline characteristics for this study are defined in **2.5** Table 2.6 and Table 2.7. Two tSNPs showed strong association with sPLA2 mass. For rs11573156 C>G in the 5'UTR of *PLA2G2A* (which is potentially part of the *PLA2G2A* promoter) (Figure 1.18), compared to the common CC genotype, the G homozygotes had 2.16ng/ml higher sPLA2-IIA levels ($P=1.9 \times 10^{-14}$) from 2.59ng/ml (CC) to 4.75ng/ml (GG). The second variant, rs3767221 A>C in the 3'UTR (Figure 1.18), showed the opposite effect with the rare homozygotes having significantly lower (1.77ng/ml) sPLA2-IIA mass ($P=2.5 \times 10^{-10}$) compared to the common A variant homozygotes from 3.94ng/ml (AA) to 2.17ng/ml (CC) (Wootton, Drenos et al. 2006).

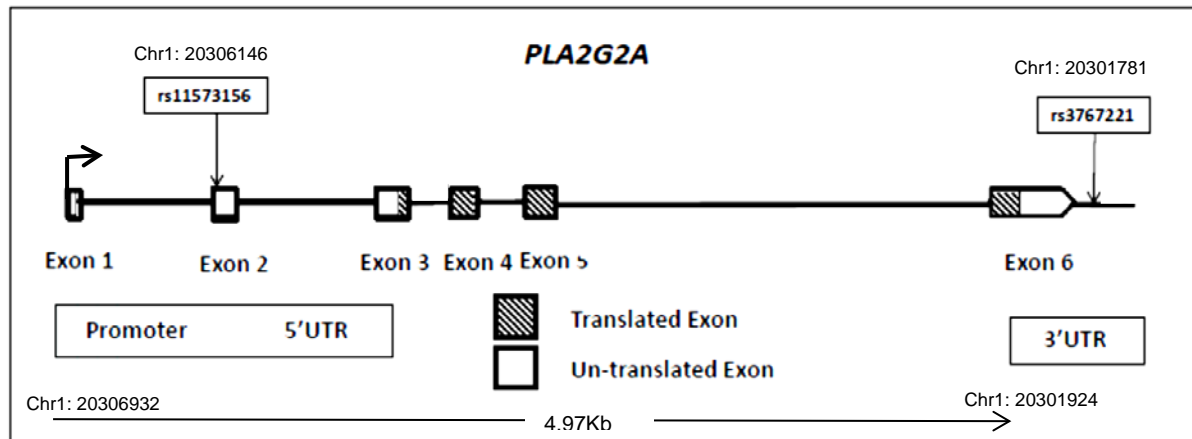


Figure 1.18. A Schematic Representation *PLA2G2A* and SNPs of Interest. This figure represents a map of the human *PLA2G2A* gene and our SNPs of interest, rs11573156 C>G and rs3767221A>C. Coding exons are shown as hatched blocks, non-coding exons are shown as empty blocks. The gene is reverse transcribed and this schematic is based on the Ensembl transcript; Transcript: ENST00000400520. This transcript relates to Gene: ENSG00000188257 and Protein: ENSP00000383364 (Based on data from Ensembl.org). Data from PharmGKB predicts that the 5'UTR, including untranslated exon1 (54bp), untranslated exon 2 (110bp), the untranslated portion of exon 3 (105bp) and the additional untranslated sequence = 1638bp. The translated portion of exon 3 = 39bp, intron 1= 208bp, exon 4 = 144bp, intron 2 = 257bp, exon 5 = 106bp, intron 3 = 2,170bp, exon 6 translated = 142bp, exon 6 untranslated = 2,66bp and the 3'UTR = 263bp.

Additionally, our lab had previously identified seven tagging SNPs (Wootton, Arora et al. 2007), spanning the gene *PLA2G5* (Figure 1.19 Table 1.5) were analysed in a cohort of Type II diabetes patients (UDACS). From these original 7 tSNPs, four SNPs were selected for further analysis on the basis of positive associations with sPLA2 mass and LDL cholesterol from univariate analysis in this study (Wootton, Arora et al. 2007).

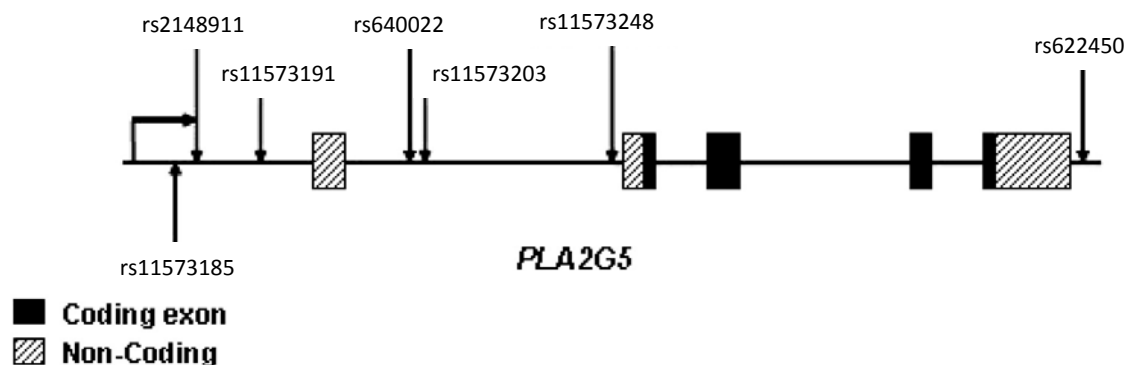


Figure 1.19. A schematic of the *PLA2G5* gene and the 7 tSNPs originally identified (Wootton, Arora et al. 2007). Schematic reprinted from (Wootton, Arora et al. 2007). The UTR regions were defined in data from the Ensembl Genome browser as; 5'UTR= Chr1:20,396,451-20,396,700 and 3'UTR= Chr1:20,417,622-20,417,911. The region spanned here is from the upstream 5'UTR SNP rs11573185 to downstream 3'UTR SNP rs622450, since these were the most extremely located previously identified tSNPs; Chr1: 20,395,401-20,419,340.

SNPs and minor allele frequency of the <i>PLA2G5</i> tSNPs in UDACS	
rs number	Minor allele frequency (95% CI)
rs11573185	0.45 (0.43-0.48)
rs2148911	0.07 (0.05-0.08)
rs11573191	0.18 (0.16-0.19)
rs640022	0.15 (0.13-0.16)
rs11573203	0.28 (0.26-0.30)
rs11573248	0.36 (0.34-0.39)
rs622450	0.14 (0.12-0.15)

Table 1.5. The seven *PLA2G5* tSNPs originally identified by Wootton et al. and their minor allele frequencies (Wootton, Arora et al. 2007).

Genetic variations in *PLA2G10* have been utilised in only a few studies to assess the potential contribution of sPLA2-X to overall sPLA2 activity. However, the increased potency of sPLA2-X compared to sPLA2-IIA and even sPLA2-V (Ishimoto, Yamada et al. 2003) suggests that it could have a more significant effect on total sPLA2 activity, dependent on sufficient expression in a tissue of interest and therefore suggests further studies using a genetic approach would be valid.

We identified 10 tagging-SNPs in *PLA2G10* using the STRAM algorithm (Burkett, Ghadessi et al. 2005). Five of these tSNPs occurred at frequencies <5% and were not

studied further and an additional three of these tSNPs were analysed in additional cohorts. An additional SNP, the C-SNP rs4003228 (R38C) had been reported by Gora et al (Gora, Perret et al. 2009), and was included in our study despite the low reported MAF (3%).

A previous study (Gora, Perret et al. 2009) analysed and identified several potential SNPs of interest within the *PLA2G10* gene in two cohorts; *SIPLAC*, a nested case control study of MI based on the trends and determinants in cardiovascular disease project registers recruited from samples in Belfast and Glasgow (n=312) and *AtheroGene*, a cohort of CAD patients recruited from the Department of Medicine II of the Johannes Gutenberg University Mainz and the Bundeswehrzentral Krankenhaus Koblenz at the occasion of a diagnostic coronary angiography (n=1,303). A priori inclusion criterion was the presence of a diameter stenosis >30% in at least one major coronary artery. Patients in this study were followed up for a median period of 6.2 years (Gora, Perret et al. 2009).

In functional work (Gora, Perret et al. 2009) rs4003228 (R38C) was identified as the SNP with the most potential to affect the contribution of sPLA2-X to sPLA2 activity. The SNP was not shown to affect the risk of CHD in the combined *AtheroGene* and *SIPLAC* data, however this was potentially due to the lower MAF of this SNP (3%). Data from the R38C functional experiments (Gora, Perret et al. 2009) suggested that R38C may play a role in controlling sPLA2-X activity. This substitution leads to an odd number of cysteines in the sPLA2-X protein, which could result in misfolding, mistargeting and subsequent protein degradation. Fluorescence immunocytochemistry experiments were carried out in COS-7 cells which were transiently transfected with both R38C wild-type and rare sPLA2-X cDNA. Both forms of the protein were shown to be expressed in the endoplasmic reticulum and Golgi compartments. Analysis of the expression patterns and activity of both forms of the protein showed that the presence of the rare allele lowered the amount of sPLA2 protein produced by COS-7 cells, preventing its efficient secretion and it had low or no activity. These results suggested that the rs4003228 may have a potential impact on sPLA2-X activity (Gora, Perret et al. 2009).

A second SNP, rs36072688 T>C, was also shown to have a potential to effect the risk of future cardiovascular events, but only in the *AtheroGene* study and not in a case-control cohort. This SNP is located in the 5'UTR of *PLA2G10* and requires further research.

1.7.3. Mendelian Randomisation

There are many potential biases that can affect observational epidemiological studies. These include reverse causation and confounding and any of these biases can interfere with study outcomes and prevent a robust identification of causal associations. There have been several examples of randomised control trials (RCT) examining an

intervention that has been assessed as potentially causal in observational and/or animal studies and yet produced a different outcome from the RCT of a drug based intervention. For example, observational data has shown the potential causality of the Hedgehog pathway in pancreatic cancer development (Thayer, di Magliano et al. 2003), leading to the development of the drug IPI-926, to be co-administered with the known pancreatic cancer drug, gemcitabine. Results from murine studies have previously indicated that co-administration of the drug IPI-926, a Hedgehog cellular signalling pathway inhibitor, with gemcitabine can help to stabilize the disease (Olive, Jacobetz et al. 2009). Randomised control trials, while being the gold standard for determining causality are expensive and time consuming. They can also be risky, since causality cannot be proven by observational and animal data alone. In the case of IPI-926 the drug company Infinity Pharmaceuticals, which owns rights to the compound, previously announced that it was ending trials of the drug for pancreatic cancer due to poor preliminary results in 122 patients with previously untreated metastatic pancreatic cancer. However, it is still unclear whether the Hedgehog pathway plays a causal role in pancreatic cancer or is a marker for the disease; is the drug ineffective or is the biomarker not causal? A method that could help to answer this question without the need for developing another molecule to trial is Mendelian randomisation.

Mendelian randomisation is a technique that takes advantage of natural genetic randomisation during conception. The random allocation of alleles at conception, lends itself to the likelihood that all other traits, besides the one you are investigating, are evenly distributed between the genotype groups, similar to the blind allocation of intervention and placebo in an RCT. Therefore the alleles are not affected by reverse causation. These germline genetic variants can act as proxy for environmentally modifiable biomarkers by acting as a genetic instrument for these biomarkers in observational studies. By separating individuals by genotype as either homozygous for common variant, homozygous for the rare allele or heterozygous and analysing changes to a biomarker and disease outcomes in each genetic group, it can be determined whether a particular genotype for a biomarker-affecting variant is more or less likely to be at risk for a disease without reverse causation. Regarding confounding, the nature of this technique is such that the observational outcomes for confounding factors are calculated at the same time as the associations for the potentially causal biomarker. If the genetic variation has no effect on any other biomarker besides the disease outcome it is strong evidence that any association with a disease can be attributed to the biomarker of interest. Mendelian randomisation can therefore be considered naturally analogous for randomised control trials (Figure 1.20), but at a fraction of the cost and time (Hingorani and Humphries 2005; Lawlor, Harbord et al. 2008).

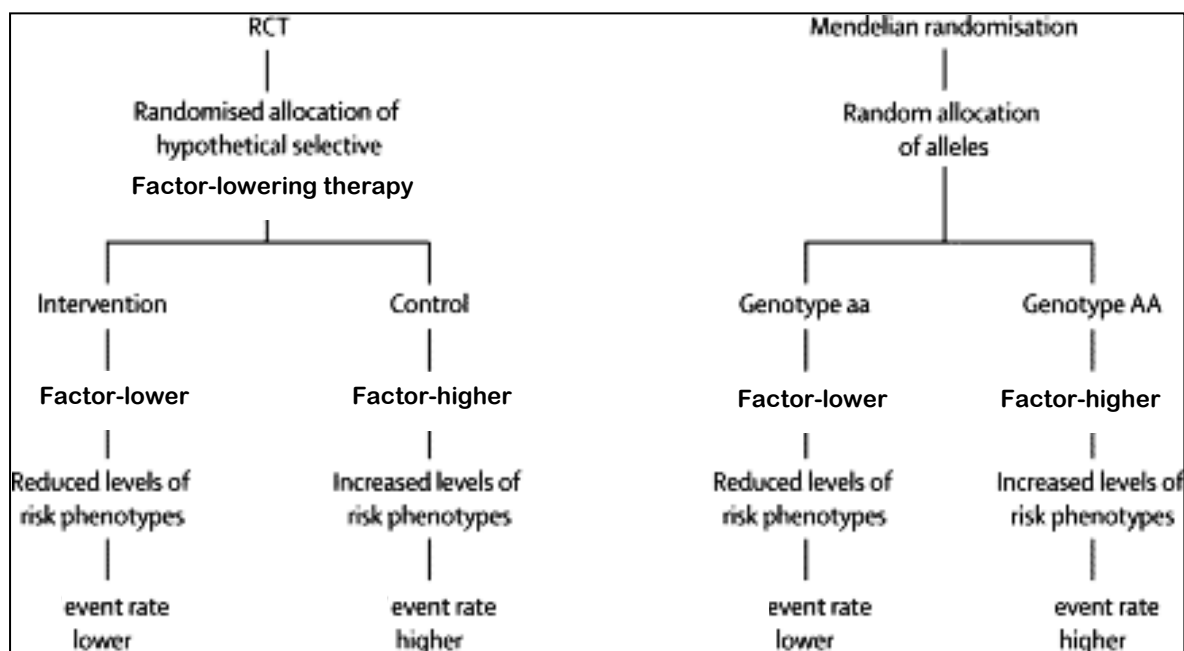


Figure 1.20. Mendelian Randomisation is Analogous for an RCT. This figure shows the parallels between a randomised control trial (RCT) and the use of Mendelian randomisation in conjunction with an epidemiological observational study. Reprinted from: (Hingorani and Humphries 2005).

An excellent example of how Mendelian randomisation has been used to great success was after the recent termination of a Phase III trial for the CETP inhibitor, torcetrapib (Barter, Caulfield et al. 2007). Torcetrapib was a first in class inhibitor of CETP, which is known to increase HDL levels. The trial was halted prematurely in 2006 due to an unexpected increase in blood pressure and cardiovascular events associated with the drug (Barter, Caulfield et al. 2007). This was despite positive results from a Phase II trial (Jensen and Hampton 2007). What was unclear was whether this increase in blood pressure was due to an off-target effect of the drug or as a direct result of CETP inhibition. It was essential to investigate this since other novel CETP inhibitors were already being produced. Should CETP inhibition raise blood pressure then it would be unethical to use any CETP inhibitor to raise HDL in patients with a high risk of CHD. However, if the increase in blood pressure could be attributed to an off-target effect of torcetrapib specifically then it would be very worthwhile to continue with research into other CETP inhibitor drugs. It would be costly and morally questionable to attempt to trial another CETP inhibitor without first knowing whether it was likely to endanger patients. By employing Mendelian randomisation to compare a group with naturally decreased CETP activity with a group with normal CETP activity, one could conclude whether lower CETP is naturally associated with an increase in blood

pressure (Sofat, Hingorani et al. 2010). If not then it would be a reasonable conclusion that it is an off target effect of torcetrapib specifically. A large Mendelian randomisation study was carried out using *CETP* genetic variants which mimicked the effects of torcetrapib by inhibiting CETP and increasing HDL levels. A total of 67,687 individuals from genetic studies and 17,911 from randomized trials were genotyped for these variants and their effects on several cardiovascular risk factors, including blood pressure, were compared (Sofat, Hingorani et al. 2010). These results were then matched with those of the torcetrapib trial. The study showed no significant differences in blood pressure associated with the genetic variant for lower CETP. This suggests that inhibition of the CETP pathway is still a valid research tool to treat those at a high risk of CHD since the increased blood pressure associated with torcetrapib appears to be an off-target of this drug only and a different class of drugs is unlikely to cause the same problem (Sofat, Hingorani et al. 2010).

Additionally an MR study was carried out with the inflammatory marker, IL-6 where IL-6 polymorphisms have been used to assess the potential of IL-6 receptor blockers as an intervention for CHD (Interleukin-6 Receptor Mendelian Randomisation Analysis, Hingorani et al. 2012). The data showed favourable results regarding IL-6R inhibition as a mechanism for CHD prevention (Interleukin-6 Receptor Mendelian Randomisation Analysis, Hingorani et al. 2012). Conversely, recent MR results suggest that C-reactive protein inhibition is not a viable intervention to prevent CHD events (Collaboration, Wensley et al. 2011). Despite showing a strong independent association with CHD in observational studies, when selected CRP polymorphisms were analysed in an MR setting it was concluded that CRP is unlikely to play a causal role in the disease (Collaboration, Wensley et al. 2011). This was supported by several additional studies (Collaboration 2008; Elliott, Chambers et al. 2009) and provides researchers with strong evidence against developing CRP inhibitors with the intention of reducing CHD.

1.8. Hypothesis

That secretory phospholipase A2 enzymes, sPLA2-IIA, sPLA2-V and sPLA2-X have a significant, independent effect on the risk of CHD.

1.9. Aims

The following specific aims will be addressed to support the hypothesis outlined above

- i. To identify common genetic polymorphisms in *PLA2G2A*, *PLA2G5* and *PLA2G10* that are associated with increased sPLA2-IIA levels and/or sPLA2 activity.
- ii. To examine the functionality of *PLA2G2A* common polymorphisms associated with sPLA2-IIA levels and activity in order to establish a biological mechanism and verify the polymorphisms' suitability as genetic instruments for Mendelian Randomisation.
- iii. To genotype the genetic variants identified in aim (i) and verified in aim (ii) in a number of large prospective and case control studies.
- iv. To use the data generated in aim (iii) for *PLA2G2A* common polymorphisms in a large Mendelian Randomisation study to examine the association of sPLA2-IIA with CHD and a number of established CHD biomarkers and its likely contribution to sPLA2 activity measures.
- v. To use the data generated in aim (iii) for both *PLA2G5* and *PLA2G10* common polymorphisms to examine their potential association to CHD and CHD biomarkers and their potential contribution to sPLA2 activity measures.

CHAPTER 2

MATERIALS AND METHODS

Please note: Gloves were used throughout for all methods. Filter tips were used and protective garments worn throughout where appropriate. All stocks and solutions were sterilised where necessary using standard conditions (121°C for 1 hour).

2.1. Commonly Used Reagents, Stocks, Products and Kits (alphabetical)

Reagent/Stock/Kit:	Supplier:	Contains:
Agarose	Invitrogen (Life technologies), Carlsbad, California, USA (Catalogue No: 16500-500)	UltraPure Agarose, 500g powder
Ammonium Persulphate (APS)	Severn Biotech (Worcester, UK), (Catalogue No: 20-3001-25)	25g Ammonium Persulphate (APS), >99% purity (228.2g/mol)
Ampicillin	Sigma-Aldrich (St. Louis, MO, USA (Catalogue No: A9518)	Ampicillin sodium salt (C ₁₆ H ₁₈ N ₃ NaO ₄ S) Ampicillin solution made up with dH ₂ O (50mg/ml)
Annealing Buffer (10x)	Made up in house	0.2M Tris (pH 7.6) 0.1M MgCl ₂ 0.5M NaCl
B-Galactosidase (β-Gal) Staining Kit	Promega (Madison, Wisconsin, USA) (Catalogue No: E1081)	β-Gal Staining plasmid pSV-β-Galactosidase Control Vector (20µg)
Binding Buffer (10x) (EMSA)	Made up in house	0.1M Tris 0.5M KCl pH 7.5
Biotin-11-dUTP	Thermo Scientific, Fermentas Molecular Biology Tools (ThermoFisher Scientific, Waltham, MA, USA), (Catalogue No: R0081)	Biotin-11-dUTP (50nM)
Buffer A (Nuclear Extract)	Made up in house	10mM HEPES, pH 7.9, 4°C 1.5mM MgCl ₂ 10mM KCl 0.5mM dithiothreitol (DTT) (C ₄ H ₁₀ O ₂ S ₂)
Buffer C (Nuclear Extract)	Made up in house	20mM HEPES, pH 7.9, 4°C 25% v/v, Glycerol, 0.42M NaCl, 1.5mM MgCl ₂ , 0.2M EDTA, 0.5mM PMSF, 0.5mM DTT
Cell Fixture Solution	Made up in house	10% Phosphate buffer solution (PBS) (10x), 5% Formalin (40% formaldehyde solution), 0.2% glutaraldehyde, dH ₂ O
Chloroform (≥99%)	Sigma-Aldrich (St. Louis, MO, USA). (Catalogue No.: C2432-500ML).	Chloroform (CHCl ₃) with amylenes as a stabilizer (Molecular weight: 119.38) (500ml)
Deoxyribonucleotide triphosphate (dNTP) Mix	New England Biolabs (Ipswich, MA, USA), (Catalogue No.: NO447S).	Deoxynucleotide Solution Mix, equimolar solution of dATP, dCTP, dGTP and dTTP (10mM per nucleotide triphosphate)
Dimethyl sulfoxide (DMSO) (10x)	Sigma-Aldrich (St. Louis, MO, USA) (Catalogue No.: D9170-5VL).	Dimethyl sulfoxide ((CH ₃) ₂ SO) PCR reagent, 5ml
DH5α Derived Competent <i>E.Coli</i> Cells	New England BioLabs (Ipswich, MA, USA) (Catalogue No.: C2988J)	NEB 5-α Competent <i>E.Coli</i> (Subcloning efficiency), 6 tubes (0.4ml/tube)
DNA Ladder	Invitrogen (Life technologies, Carlsbad, California, USA), (Catalogue No.: 10787-018).	Plus DNA Ladder (1kb), 250µg

Dulbecco's Modified Eagle Medium (DMEM) (Serum containing)	PAA The Cell Culture Company (GE Healthcare, Pasching, Austria) (Catalogue No.: E15-810), PAA The Cell Culture Company (GE Healthcare, Pasching, Austria), (Catalogue No.: A15-151)	DMEM with 4.5g/L glucose and L-Glutamine 10% Foetal bovine serum (Gold)
Electrophoretic Mobility Shift Assay (EMSA) Detection Kit	Thermo Scientific (ThermoFisher Scientific, Waltham, MA, USA), (Catalogue No.: 20148)	LightShift Chemiluminescent EMSA Kit (100 reactions per kit)
Ethidium Bromide	Invitrogen (Life technologies, Carlsbad, California, USA), (Catalogue No.: 15585-011)	UltraPure Ethidium Bromide solution (10mg/ml)
Glycerol	Made up in house	50% v/v Glycerol (C ₃ H ₈ O ₃) dH ₂ O
Human hepatocellular carcinoma (Huh 7) Cell Line	Health Protection Agency Culture Collections, (Catalogue No.: 01042712)	Human hepatocellular carcinoma (Huh-7D12) cell line
Hybond-N+ Membrane	GE Healthcare (Thermo Fisher Scientific, Waltham, MA, USA). (Catalogue No.: GZRPN3050B).	Hybond-N+ (300mmX3M)
Isoamyl Alcohol (≥98%)	Sigma-Aldrich (St.Louis, MO, USA), (Catalogue No.: W205702-1KG-K).	Isoamyl (CH ₃) ₂ CHCH ₂ CH ₂ OH) (Molecular weight: 88.15)
Lipofectamine™ 2000	Invitrogen (Life technologies, Carlsbad, California, USA), (Catalogue No.: 11668-019)	1.5ml Lipofectamine™ 2000 Transfection Reagent (1mg/ml)
Loading Buffer (6x) (Agarose gels)	Made up in house	10mM Tris-HCl (pH 7.6), 0.03% Bromophenol Blue, 0.03% Xylene Cyanol CFF, 60% glycerol (C ₃ H ₈ O ₃), 60mM, Ethylenediaminetetraacetic (EDTA) acid
Loading Buffer (EMSA)	Made up in house	50% v/v: 10x Xylene Cyanol CFF loading buffer, 10x Bromophenol Blue loading buffer
Luciferase Assay Kit	Made up in house	50% v/v: 10x Xylene Cyanol CFF loading buffer, 10x Bromophenol Blue loading buffer
Luria Bertani (LB)-Agar	Made up in house	Per 100ml: 1g Tryptone (C ₄ H ₁₁ NO ₃ -HCl), 0.5g yeast extract, 1g sodium chloride (NaCl), 0.1ml of 5M sodium hydroxide (NaOH), 1.5g Agar, 99.9ml dH ₂ O
Luria Bertani (LB)-Medium	Made up in house	Per 100ml: 1g Tryptone (C ₄ H ₁₁ NO ₃ -HCl), 0.5g sodium chloride (NaCl), 0.5g yeast extract, 0.2g glucose (C ₆ H ₁₂ O ₆), 100ml dH ₂ O
Lymphocyte Separation Solution	PAA: The Cell Culture Company (GE Healthcare, Pasching, Austria). (Catalogue No.: J11-004).	Lymphocyte Separation Medium LSM 1077 with Ficoll™ density gradient media (1,077g/ml)
Magnesium Chloride (MgCl ₂)	See Taq DNA Polymerase (Recombinant) . MgCl ₂ is provided with this kit (Invitrogen (Life technologies, Carlsbad, California, USA), Catalogue No.: 10342-020)	50mM MgCl ₂
Maxiprep Kit	Sigma-Aldrich (St. Louis, MO, USA), (Catalogue No.: NA0310-1KT).	GenElute™ HP Plasmid Maxiprep Kit, used to isolate plasmid DNA up to 1.2mg (25x reaction kit)
Micro Array Diagonal Gel Electrophoresis (MADGE) Gel (6%)	Made up in house	For 1 x 1.5mm gel: 12ml 30% polyacrylamide (37:5:1), 47ml dH ₂ O, 3ml TBE (10x), 50µl TEMED, 500µl Ammonium Persulphate (APS) 10x)
Miniprep Kit	Qiagen (Hilden, Germany), (Catalogue No.: 27104).	QIAprep Spin Miniprep Kit, for purification of up to 20µg molecular biology grade plasmid DNA (50x reaction kit)
NH ₄ Polmix Buffer (10x)	Made up in house	16mM [NH ₄] ₂ SO ₄ 67mM Tris-HCL, pH8.4 0.01% Tween 20 2mM dATP, 2mM dTTP, 2mM dGTP & 2mM dCTP

Opti-Mem®	Gibco (Life technologies, Carlsbad, California, USA), (Catalogue No.: 31985-062)	Opti-Mem (Opti-Minimal Essential Medium) (100ml)
pGL3 Luciferase Reporter Vectors	Promega (Madison, Wisconsin, USA), pGL3 (Catalogue No.: E1751) (Catalogue No.: E1761) (Catalogue No.: E1741)	Luciferase Reporter Vectors for quantitative analysis of potential regulatory factors of mammalian genes. pGL3-Basic vector (20µg) (Figure 2.1), pGL3-Promoter vector (20µg) (Figure 2.2), pGL3-Control vector (20µg), (Figure 2.3)
Phosphate Buffer Saline (PBS)	Gibco, (Catalogue No.: 18912014)	100 solid tablets, each to be dissolved in 500ml dH ₂ O (1x)
Polyacrylamide	Severn Biotech (Worcester, UK) (Catalogue No.: 20-2300-10) (Catalogue No.: 20-2100-10)	(30%) (19:1) 30% Acrylamide Gel solution Bis-Acrylamide (w/v) Ratio 19:1 (concentrate) (30%) (37:5:1) 30% Acrylamide concentrate Bis-Acrylamide (w/v) Ratio 37:5:1 Protein Gel
Primers	Designed using the program Primer3 version 4.0 (http://frodo.wi.mit.edu/primer3/). Primers synthesised by Eurofins MWG Operon (Ebersberg, Germany)	All primers diluted to 100pmol/µl as standard.
pRL-TK <i>Renilla</i> Luciferase Control Reporter Assay	Promega (Madison, Wisconsin, USA), (Catalogue No.: E2241)	pRL-TK Vector (20µg). An internal control reporter used in combination with an experimental reporter vector (e.g. pGL3 reporter vectors) to cotransfect mammalian cells. Used with the Dual Luciferase Reporter Assay
Protease Inhibitor	Thermo Scientific, Pierce Protein Research Products (ThermoFisher Scientific, Waltham, MA, USA) (Catalogue No.: 78410)	2ml Protease Inhibitor Cocktail Kit (100x)
pUC19	New England Biolabs (Ipswich, MA, USA), (Catalogue No.: N3041L).	pUC19, a commonly used plasmid cloning vector in <i>E.coli</i> (1,000µg/ml)
Random Primers (d(N) ₆)	Invitrogen, (Life technologies, Carlsbad, California, USA), (Catalogue No.: 48190-011)	9 Units Random Primers (Hexamers) (3µg/µl in 3mM Tris-HCl (pH 7.0), 0.2mM EDTA)
Restriction Enzymes and buffers	New England Biolabs (Ipswich, MA, USA) (Individual enzyme catalogue Numbers can be found at: http://www.neb.com/nebecomm/products/category1.asp)	Known Restriction Endonucleases with relative consensus sequences and corresponding NEB Buffers (10x)
Ribonuclease (RNase) H	New England BioLabs (Ipswich, MA), (Catalogue No.: M0297S)	250 units Ribonuclease H (RNase H) (5000U/ml)
RNeasy Mini Kit	Qiagen (Hilden, Germany). (Catalogue No.: 74104).	RNeasy Mini Kit, for purification of up to 100µg total RNA from cells, tissues and Yeast (50 reaction kit)
Saline Sodium Citrate (SSC) Buffer (20x)	Made up in house	3M sodium chloride (NaCl) 0.3M sodium citrate dH ₂ O
Site Directed Mutagenesis Kit	Stratagene (California, USA), QuikChange® (Catalogue No.: 210518)	Lightning Site-Directed Mutagenesis Kit
Superscript III Reverse Transcription Kit	Invitrogen, (Life technologies, Carlsbad, California, USA) (Catalogue No.: 18080-044)	10,000 Units Superscript III Reverse Transcriptase (200U/µl) First Stand Buffer (5x) Dithiothreitol (DTT) 100mM
T4 DNA Ligase	New England BioLabs (Ipswich, MA, USA), (Catalogue No.: M0202S)	T4 DNA Ligase kit
Taq DNA Polymerase (Recombinant)	Invitrogen (Life technologies, Carlsbad, California, USA) (Catalogue No.: 10342-020)	Thermostable <i>Thermus aquaticus</i> (Taq) DNA Polymerase (500 units at 5 U/µl). Amplifies products up to 5Kb. 50mM MgCl ₂
TaqMan® SNP Genotyping Assays	Applied Biosystems (Life technologies, Carlsbad, California, USA), (Part No.: 4351379). Table 2.2. Individual TaqMan® SNP Genotyping Assay List	Human Pre-Designed SNP Genotyping Assays (40x), small scale (Table 2.2)
TaqMan® (Genotyping) Universal PCR Master Mix	Applied Biosystems (Life technologies, Carlsbad, California, USA), (Part No.: 4371355)	TaqMan Genotyping Master Mix (2x) with AmpliTaq Gold DNA polymerase, 1- Pack (1x 10ml)

TaqMan® Gene Expression Assays & Endogenous Controls	Applied Biosystems (Life technologies, Carlsbad, California, USA), (Part No.: 4331182). Table 2.3. Individual TaqMan® Gene Expression Assay List	Small Inventoried TaqMan Gene Expression or Endogenous Control Assay (20x) (Table 2.3)
TaqMan® Gene Expression Master Mix	Applied Biosystems (Life technologies, Carlsbad, California, USA), (Part No.: 4369016)	TaqMan Gene Expression Master Mix (2x), 1-Pack (1x 5ml)
Terminal Deoxynucleotidyl Transferase (TdT)	Thermo Scientific, Fermentas Molecular Biology Tools (ThermoFisher Scientific, Waltham, MA, USA) (Catalogue No.: EP0161)	500 Units Terminal Deoxynucleotidyl Transferase (TdT) enzyme (20U/μl) 0.4ml TdT Reaction Buffer (5x)
Tetramethylethylenediamine (TEMED)	VWR International (Radnor, Pennsylvania, USA), (Catalogue No.: 443083G)	TEMED (NNN'N'-Tetramethylethylenediamine) Electron, (CH ₃) ₂ NCH ₂ CH ₂ N(CH ₃) ₂ , (116.21g/mol)
Tris-Borate-EDTA (TBE) Buffer (10x)	Fisher BioReagents, (product No.: BP13334)	4L PolyPac
Trypsin	Gibco (Life technologies, Carlsbad, California, USA), (Catalogue No.: 25200-056)	0.25% Trypsin-EDTA (1x), phenol red
Whatman Chromatography Paper	Whatman (GE Healthcare) From Fisher Scientific (Thermo Fisher Scientific, Waltham, MA, USA), (Product Code: CJF-240-070E)	Whatman chromatography paper No 3MM Chr electrophoresis and chromatography thick smooth surfaced 75mm, width 100m
White Cell Culture Plates	Appleton Woods (Appleton Woods Limited, Birmingham, UK), (Code: CC701, Alt Code: 3362)	96 well plate, flat-well, sterile, with lid, Corning (100 pack)
X- Galactosidase (X-gal)	Made up in house	10% Phosphate buffer solution (PBS) (10x), 10% Potassium ferricyanide (50mM), 10% Potassium ferrocyanide (50mM), 0.2% MgCl ₂ (1M), 2% X-Gal (20mg/ml)

Table 2.1. Is a comprehensive list of commonly used reagents, stocks and kits. It includes details of the production company and what each kit contains or what each reagent/stock is made up from.

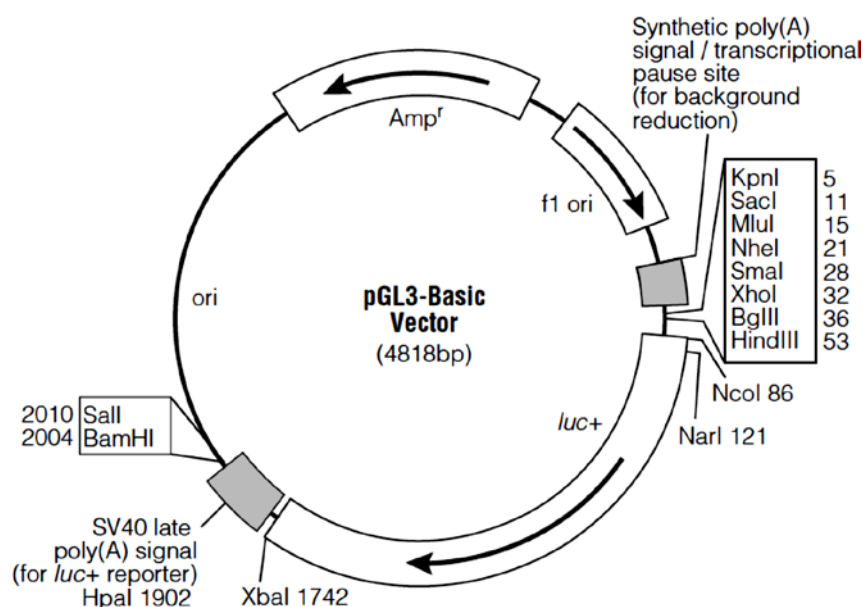


Figure 2.1. The pGL3-Basic Luciferase Reporter Vector. This vector includes, 2 polylinker sites, the firefly and *renilla* luciferase (*luc+*) genes, a SV40 late poly(A) signal and the Amp^r, ampicillin resistance region.

([www.promega.com/~media/Files/Resources/Protocols/Technical Manuals/0/pGL3 Luciferase Reporter Vectors Protocol.ashx](http://www.promega.com/~media/Files/Resources/Protocols/Technical%20Manuals/0/pGL3%20Luciferase%20Reporter%20Vectors%20Protocol.ashx))

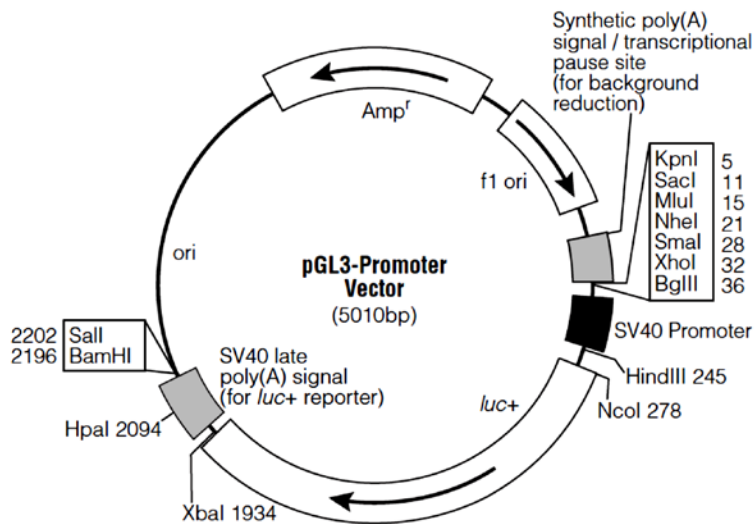


Figure 2.2. The pGL3-Promoter Luciferase Reporter Vector. This vector includes, 2 polylinker sites an SV40 promoter, the firefly and *renilla* luciferase (*luc+*) genes, the SV40 late poly(A) signal and the Amp^r, ampicillin resistance region. ([www.promega.com/~media/Files/Resources/Protocols/Technical Manuals/0/pGL3 Luciferase Reporter Vectors Protocol.ashx](http://www.promega.com/~media/Files/Resources/Protocols/Technical%20Manuals/0/pGL3%20Luciferase%20Reporter%20Vectors%20Protocol.ashx))

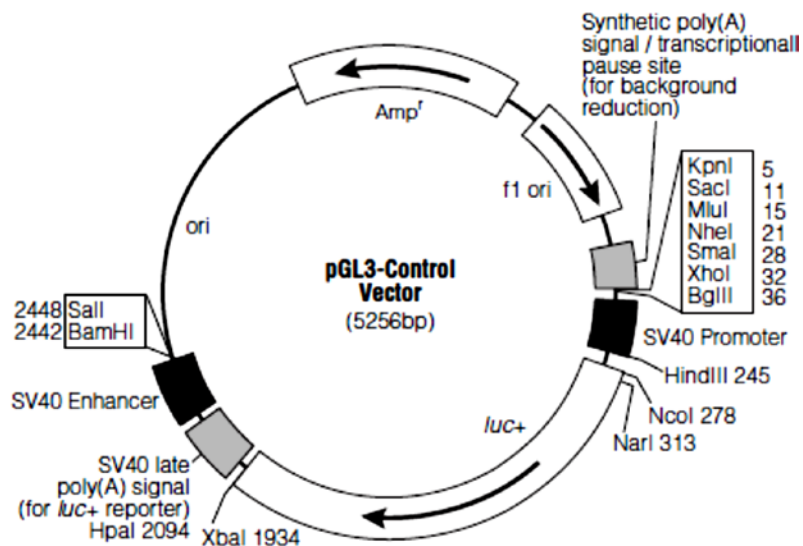


Figure 2.3. pGL3-Control Vector. Contains an SV40 promoter and enhancer sequences that result in a strong expression of both the firefly and *renilla* luciferase (*luc+*) genes in many mammalian cell types. This plasmid is often used as a general check for transfection efficiency and as an internal standard for promoter and enhancer enzyme activities expressed by pGL3 recombinants ([www.promega.com/~media/Files/Resources/Protocols/Technical Manuals/0/pGL3 Luciferase Reporter Vectors Protocol.ashx](http://www.promega.com/~media/Files/Resources/Protocols/Technical%20Manuals/0/pGL3%20Luciferase%20Reporter%20Vectors%20Protocol.ashx))

Gene	SNP	Assay Number	Vic Probe Genotype	Fam probe Genotype
PLA2G2A	rs11573156	PLA2C763G	C	G
PLA2G2A	rs3767221	PLA2T5128G	T	G
PLA2G5	rs11573185	PLA2G5C100A	C	A
PLA2G5	rs11573248	P5TAD12605	TA	-
PLA2G5	rs11573191	PLG5G1114A	G	A
PLA2G5	rs11573203	PG5GDE3272	G	-
PLA2G10	rs72546339	PLG10C264A	C	A
PLA2G10	rs72546340	PLG10G303C	G	C

Table 2.2. Individual TaqMan® SNP Genotyping Assay List

Gene	Target or Endogenous Control	Exons Spanned	Assay Number	Reporter Dye
PLA2G2A	Target	1-2	Hs01044022_m1	FAM
PLA2G2A	Target	5-6	Hs00179898_m1	FAM
ACTB	Control	N/A	Hs99999903_m1	FAM
GAPDH	Control	N/A	Hs99999905_m1	FAM

Table 2.3. Individual TaqMan® Gene Expression Assay List

2.2. Commonly Used Websites, Software and Services (alphabetical)

Name:	Company Details:	Product Details:
1000 Genomes Browser	http://www.1000genomes.org/ensembl-browser	1000 Genomes 2008-2012 database. A deep catalogue of human genetic variation.
Anthera Pharmaceuticals	Anthera Pharmaceuticals (Hayward, CA. USA) http://www.anthera.com/	This company was developing varespladib methyl, a specific sPLA2 inhibitor.
CLUSTALW	Multiple Sequence Alignment by CLUSTALW (http://www.genome.jp/tools/clustalw/)	Compares two FASTA sequences for alignment. It is often used to check the sequence of a DNA product compared to a known consensus sequence.
Ensembl Genome Browser	www.ensembl.org .	International genome database project run by the Ensembl Scientific Advisory Board
FinchTV	PerkinElmer (Geospiza, Seattle, WA, USA)	FinchTV Version 1.4 trace viewer software reads results from Source BioScience sequencing (See Sequencing DNA) and presents them as an electropherogram. This software also allows results to be read as a FASTA sequence in both forward and reverse orientations.
Genesnap	Syngene (Syngene UK, Cambridge UK) Genesnap software v6.04	Software used in conjunction with Syngene Gel Documentation System to visualise and photograph gel electrophoresis products.
International HapMap Project	A partnership of scientists and funding agencies from Canada, China, Japan, Nigeria, the United Kingdom and the United States (http://www.hapmap.ncbi.nlm.nih.gov/).	A public resource that helps to find genes associated with human disease and response to pharmaceuticals.

MirBase	MirBase The microRNA Database. Manchester University, Manchester UK. www.mirbase.org	Algorithms from this website were used to predict miRNA binding in sequences of interest.
MiRanda version 4.0	The European Bioinformatics Institute (EBI) (Wellcome Trust Genome Campus, Cambridgeshire, England) http://www.ebi.ac.uk/	Algorithms from this website were used to assess potential microRNA binding to sequences of interest.
mirTar	MicroRNA Target prediction (mirTar) (Institute of Bioinformatics National Chiao Tung University, Hsinchu, Taiwan).	A tool to easily identify the biological functions and regulatory relationships between a group of known/putative miRNAs and protein coding genes and perspective on the information about miRNA targets on alternatively spliced transcripts.
MatInspector	Genomatix Software GmbH (Munich, Germany) www.genomatix.de/online_Help/help_matinspector/matinspector_Help.html	MatInspector. Database for transcription factor prediction.
NEBCutter	New England Biolabs (Ipswich, MA, USA). NEBCutter, version 2.0, http://www.neb.com/tools/index.aspx?req=nebcutter	A tool used to identify natural cutting sites for Restriction Enzyme cutting in a sequence of interest (Vincze, T., Posfai, J. and Roberts, R.J. NEBcutter: a program to cleave DNA with restriction enzymes Nucleic Acids Res. 31: 3688-3691 (2003)).
PharmGKB	Pharmacogenomics. Knowledge. Implementation. (Stanford University USA) www.pharmkb.org	Used to identify intron/exon boundaries and 5'UTR/3'UTR boundaries.
PITA	Segal Lab of Computational Biology (Weizmann Institute, Rehovot, 76100, Israel) (http://genie.weizmann.ac.il/pubs/mir07/mir07_prediction.html).	MicroRNA target prediction algorithm.
RegRNA	A R egulatory R NA Motifs and Elements Finder (National Chiao Tung University, Taiwan) (http://regrna.mbc.nctu.edu.tw/).	MicroRNA target prediction algorithm.
REST-384 Software	Relative Expression Software Tool (REST)-384 (REST.gene-quantification.info). (Download at: (http://www.gene-quantification.de/download.html)).	Software designed to interpret gene expression data to compare genes and genotypes
Restriction Enzyme Consensus Sequences	New England Biolabs (Ipswich, MA, USA) (http://www.neb.com/nebecomm/EnzymeFinderSearchByRecombinant.asp).	A catalogue of known restriction enzyme consensus sequences
Sequence Detection System (SDS) v2.1	Applied Biosystems (Life technologies, Carlsbad, California, USA)	Sequence Detection System software version 2.4.1. High-throughput gene expression and genotyping analysis software for use with the 7900HT Fast Real-Time PCR System
Sequencing DNA	Where DNA sequences were to be verified, sequencing was carried out by Source BioScience, (London, UK)	≥10µl DNA product (100ng/µl) was sent along with known primers (3.3pmol/µl) to be sequenced. Results are supplied as a file for use with FinchTV (See FinchTV) to be viewed as an electropherogram.
SNAP Pairwise LD version 2.2.	SNP Annotation and Proxy Search (BROAD Institute) http://www.broadinstitute.org/mpg/snap/ldsearchpw.php	Input SNPs to calculate the LD between them using genotyping data from the 1000 Genomes Database and the International HapMap Project
Tm Calculator	Thermo Fisher Scientific, Waltham, MA, USA http://www.finnzymes.fi/tm_determination.html	Finnzymes PCR & qPCR-products. Calculates optimum melting temperatures.
UCSC Genome Browser	The UCSC Genome Bioinformatics Group of UC Santa Cruz, California, USA http://genome.ucsc.edu/cgi-bin/hgGateway .	Genome Reference Consortium
WinGlow Software	See Tropix TR717 Microplate Luminometer	See Tropix TR717 Microplate Luminometer

Table 2.4. A comprehensive overview of the software and bioinformatics databases used throughout this thesis. Includes company information and a brief description of use for each website.

2.3. Commonly Used Laboratory Machines and equipment (alphabetical)

Equipment:	Company:	Description:
7900HT Fast Real-Time PCR System	Applied Biosystems (Life technologies, Carlsbad, California, USA) (Catalogue No.: 4329001)	7900HT Fast Real-Time PCR System with 384 Well Block Module
Advanced Detection and Accurate Measurement (ADAM) Cell Counter	Digital Bio Pharm (Cambridge, UK) (Catalogue No.: Adam-MC)	Advanced Detection and Accurate Measurement (ADAM) Cell Counter
Biomek® 2000 Laboratory Automation Workstation	Beckman Coulter (High Wycombe, UK)	Biomek 2000 Laboratory Automation Workstation and software
Centrifuge (Up to 50ml)	Qiagen (Hilden, Germany), Centrifuge 4-16 (Catalogue No.: 81325)	Universal laboratory centrifuge with brushless rotor
Centrifuge (Up to 200ml)	Sorvall (DJB Labcare Ltd, Buckinghamshire, UK)	Sorvall RC5 centrifuge with rotors SA-600 and SA-1500 (Catalogue No. 74505)
Finnpipette multichannel dispenser	Life Sciences, Basingstoke, UK	Finnpipette multichannel dispenser
G-Strom Thermocycler	G-Strom (Somerset, UK)	G-Strom GS4 thermocycler, 4x 96 well temperature gradient blocks with optional heated lids, temperature range from 4°C-99°C
Microcentrifuge	Technico (ThermoFisher, Waltham, MA, USA) (Catalogue No.: TEC-150-030N)	Table top TEC-150-030N-Microcentrifuge Maxi; Including 18x 1.5/2ml rotor, pulse button, timer and continuous operation (Speed 1,000-14,000rpm (16,000g))
Single Channel Pipettes (P10, P20, P250, P1000)	Gilson (Gilson Inc Middleton, WI, USA)	Single channel pipettes sizes, P10, P20, P200 and P1000
Spectrophotometer (NanoDrop 8000)	Thermo Scientific (ThermoFisher Scientific, Waltham, MA, USA)	Microarray, Proteins & Labels (Catalogue No.: ND8000)
Syngene Gel Documentation System	Syngene UK (Cambridge UK),	Allows for the visualisation and photographing of gel products after electrophoresis; Used in conjunction with Genesnap v604 software
Tropix TR717 Microplate Luminometer	Applied Biosystems (Life technologies, Carlsbad, California, USA)	Tropix TR717 Microplate Luminometer; Includes WinGlow software for measurement and analysis of luciferase activity
UV Stratalinker 2400	Stratagene (Stratagene, California, USA) (Catalogue No.: 400075 (120V))	UV Stratalinker 2400; Used for membrane blotting, DNA nicking, Gene Mapping and RecA Screening

Table 2.5. Displays a comprehensive overview of all instruments used throughout this thesis. Include company data and a brief overview of what each instrument is used for.

2.4. Standard Methods

2.4.1. SNP Selection and Bioinformatics

Tagging SNPs for each *PLA2* gene were identified as described in the **Introduction (1.7.2)**. European populations were selected for without exception across all databases. Table 2.4 lists all of these bioinformatics tools, what they were used for and the web link to each one. The general parameters used per tool to research each gene/SNP are as follows:

International HapMap Project: Choose the browser release #28 including phase 1, 2 and 3. Type gene name or SNP rs number into the search box. Standard search parameters were kept and the 'Download Decorated FASTA File' Report & Analysis function was selected to view results.

MatInspector: Consensus FASTA sequences were selected from the Ensembl Genome Browser and manually copied into the MatInspector tool. The software uses a large library of matrix descriptions for transcription factor binding sites, locating matches within the FASTA sequence and assigns a quality rating to each match. I selected each FASTA sequence to include either the common or rare variants of a SNP of interest for comparison to assess whether there was a possibility of differential TF binding between the alleles. The parameters selected for were as follows: Sequence added manually using 'Your Sequence' function. *Homo Sapiens* option selected for 'search corresponding promoters'. 'Transcription Factor Binding (weight matrices)' selected for the library option.

UCSC Genome Browser Bioinformatics Browser: From the main menu. The 'Assembly' chosen was Mar. 2006 (NCBI36/hg18). The appropriate SNP or gene was selected for and the standard database parameters were applied on selection. These included Encode transcription factor binding prediction, DNA hypersensitivity prediction, spliced human expressed sequence tags (splice sites) and known SNPs.

PITA, RegRNA and MiRTar were used to predict potential microRNA (miRNA) binding sites in the *PLA2G2A* 5'UTR. An mRNA consensus sequence was entered into each target prediction algorithm that includes searches outside of the 3'UTR. The PITA and miRbase algorithms were used to predict miRNA binding in the *PLA2G2A* 3'UTR. All predictions of putative miRNA binding sites were kindly carried out under the supervision of my colleague Dr Anastasia Kalea. The specific parameters used for each prediction are quoted where appropriate.

Analysing the sPLA2 Inhibitor, Varespladib

Adhering to PRISMA guidance (Weber and Noels 2011), I conducted a systematic review of randomised trials that evaluated the effects of varespladib on sPLA2-IIA mass and other cardiovascular traits. The PubMed search parameters consisted of ("varespladib"[Substance Name] OR "varespladib methyl "[Substance Name] OR varespladib OR A002 OR LY333013). Studies in humans that were randomized trials comparing varespladib to placebo were selected for inclusion. Data was extracted and cross-checked by Dr Michael Holmes for agreement of its inclusion; discrepancies were resolved by consensus. Where possible, we harmonised variable units between studies. In the case where a conversion factor could not be identified, the standardized mean difference was used.

2.4.2. DNA Standardisation and Determining DNA Concentration

All DNA used throughout this project had been previously standardised to either: 2.5ng/μl, 5ng/μl, or 15ng/μl arrays. DNA concentration was determined throughout this thesis using the spectrophotometer, Nanodrop 8000.

2.4.3. Primer Design

Oligonucleotide primers for were designed using the Primer3 primer design program v 0.4.0 (Table 2.4), which is freely available. The design of suitable primers for a sequence of interest is based on a number of factors including; the primer melting temperature, the GC content, 3'end stability, length, secondary structure and the ability for primer-dimer formation, which is undesirable. Primers that amplified poorly were redesigned.

For the Primer3 input: A sequence including the target was identified and isolated from the Ensembl Genome Browser and pasted into the 'paste sequence' box. 'Pick left primer' and 'pick right primer' options were selected. A sequence ID was chosen to recognise what purpose the primers were for. The target sequence was identified by indicating relevant anchor base number and any surrounding base pairs that make up the target sequence. 'Excluded regions' was left blank. All other parameters were kept as the parameters common to the original template.

2.4.4. Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) is used to amplify DNA fragments of interest for an assortment of analytical techniques including genotyping, sequencing and cloning.

5-15ng of double stranded sample DNA was taken from a 96 well standardised array and was used in a standard 96 well PCR plate. DNA was transferred from array to plate using a Finnpiptette multichannel dispenser. The plate was centrifuged at 1000rpm for 1 minute to ensure that DNA was settled in the bottom of each well.

A PCR reaction mix included MgCl_2 (50mM), NH_4^+ buffer (10x), *Taq* DNA polymerase and a forward and reverse orientation primer (custom designed for the sequence of interest, usually to be around 20-30 base pairs). Once made up, this reaction mix was added to each DNA sample and a coat of 20 μ l mineral oil was added to each well to prevent evaporation. PCRs throughout this project were run on the G-Storm Thermocycler. Each reaction was optimised using test DNA samples across a gradient of MgCl_2 and temperatures.

A standard PCR has an initiation step then between 30-40 cycles of a denaturing step, an annealing step and an elongation step. The reaction concludes with a final step to ensure that all products are fully formed. The annealing temperature is based on the primer and product length. An accurate prediction of the annealing temperature for a primer is the melting temperature (T_m) minus 5. The T_m is based on the GC content and length of each primer. The annealing temperature can be calculated using the Finnzymes T_m Calculator. The timing of the elongation step is based on the length of the primers; approximately 1Kb/minute (Lusis 2000). Each cycle results in an approximately two-fold replication, accumulating a large quantity of specific DNA. Should this method fail to provide a clean amplification, 5% dimethyl sulfoxide (DMSO) was added to act as a detergent to facilitate separation of double-stranded DNA and preventing non-specific primer binding. Specific PCR reaction volumes and steps are described throughout this chapter.

2.4.5. Agarose Gel Electrophoresis

A standard 2% gel was effective for showing products in the range of 80-200bp. In order to visualize products of sizes outside this range the percentage of the gel was adjusted. A 1% gel was used to see products 500bp-1kb and a 0.5% gel to see 1-30kb.

For a standard 2% gel per 100ml: 2g agarose (Table 2.1), 10ml TBE (1x) + 0.01 μ g Ethidium Bromide (EtBr) (Table 2.1), 90ml dH_2O . The agarose and water were added together and heated in a microwave on high power for 2 minutes. Then the TBE/EtBr was added to the mix which was immediately poured into a pre-assembled mould. A comb was added to form the wells.

Usually 2 μ l of agarose gel loading buffer (6x) was added to 10 μ l of DNA product. 3 μ l of Plus DNA Ladder (1kb) (Table 2.1) was loaded into one well of each gel in order to

determine fragment size. Typical gel electrophoresis was carried out at 100V for 45 minutes, or as long as it took for the loading buffer to move down two thirds of the gel. Gels were analysed and photographed under UV illumination using the Syngene Gel Documentation System and Genesnap v6.04 software (Table 2.5). Standard parameters were maintained for image viewing with the exception of the image brightness and zoom, which were adjusted to optimum per gel.

2.4.6. Genotyping

2.4.6.1. TaqMan Genotyping

Each ABI TaqMan Genotyping Assay (40x) contains two target SNP specific primers to amplify the DNA sequence of interest (the 'DNA template') and two allele specific TaqMan MGB probes. Each allele specific probe has three main elements, a reporter dye at the 5' end and a minor groove binder (MGB) and a non-fluorescent quencher (NFQ) at the 3' end. There are two commonly used reporter dyes, VIC and FAM, and each allele is distinguished dependent on the reporter dye. The VIC dye is linked to the allele 1 probe and the FAM dye is linked to the allele 2 probe. The MGB is added in order to increase the melting temperature for a given probe length (Davies and Woolf 1993; Ross 1999; Tegos, Kalodiki et al. 2001), allowing shorter probes with an increased difference in melting temperature between matched and mismatched probes. Thus the addition of an MGB increases the efficiency of allelic discrimination. The NFQ does not fluoresce and will quench the fluorescent signal of any reporter dye linked to the 5' end for as long as the probe remains intact.

ABI TaqMan Genotyping reactions were carried out in a 384 well plate format. DNA at 5ng was transferred from a standardized array, using the Biomek 2000 Laboratory Automation Workstation (Table 2.5), into a 384 well plate and centrifuged at 3000rpm for 1 minute before being air dried. Each plate contained at least two no template control (NTC) wells.

The TaqMan genotyping reaction was carried out as per the TaqMan Genotyping mastermix manufacturers' instructions. Each plate was centrifuged at 3000rpm for 1 minute and PCR reactions were carried out using the 7900HT Fast Real-Time PCR system (Table 2.5). The standard PCR reaction conditions were as follows:

AmpliTaq Gold Enzyme Activation 95°C 10 minutes followed by 40 cycles of: Denature 92°C 15 seconds and anneal/extend 60°C 1 minute

There were several phases that occurred during each PCR reaction (Figure 2.4); Each TaqMan MGB probe annealed to the specific complimentary sequence between the forward and reverse primer sites. At this point the probe was still intact and therefore the NFQ was still at close enough proximity to the reporter dye to quench its fluorescent signal. Next the AmpliTaq Gold DNA polymerase (a component of the TaqMan Universal Master Mix) extended the DNA template between the forward and reverse bound primers and caused the cleavage of the probe, specifically by cleaving only the probes that are hybridized to the target. The effect of cleavage was to separate the reporter dye from the NFQ, which caused an increase in fluorescence from the reporter dye. Generation of increased fluorescent signal by PCR amplification indicated which alleles were present in each the sample, with a significant increase in VIC dye signal correlated with homozygosity for allele 1, an increase in FAM dye signal correlated with homozygosity for allele 2 and a significant increase in both VIC and FAM signals correlated to Alleles 1 and 2 heterozygosity (http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_039282.pdf).

TaqMan Genotyping Signal Detection

Reaction runs and plate analysis were performed using the 7900HT Fast Real-Time PCR System (Table 2.5) with Sequence Detection System (SDS) v2.1 software (Table 2.4). Standard allelic discrimination was carried out using the SDS software which included an automatic threshold signal detector. A new document was opened for each 384 well plate.

The standard allelic discrimination program assigned genotypes automatically to a plate coordinate in a text output file and in the form of an allelic discrimination plot based on the fluorescent signal detected. The light source from the cyclor excites the reporter dyes by fluorescence resonance energy transfer. VIC and FAM have significantly different excitation and emission wavelengths, which allows them to be differentially detected. Each dot on the scatterplot represents one well of the plate (either a DNA sample or an NTC). An example of the scatterplot form of results is shown below (Figure 2.5). Undefinable samples are represented by an 'X'. These samples did not reach the threshold fluorescent signal level to be defined as a particular genotype.

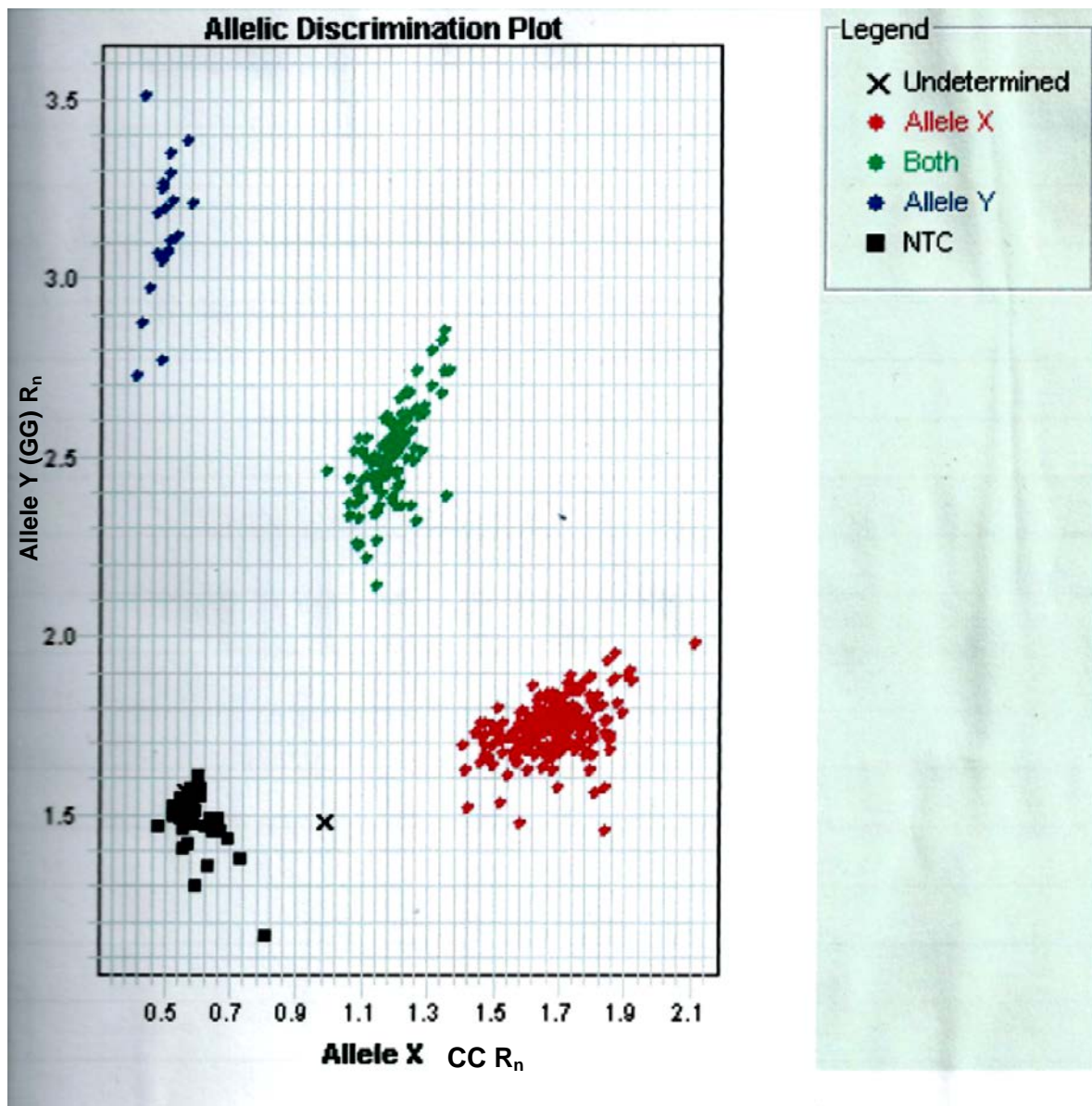


Figure 2.5. Allelic Discrimination plot. This plot represents the results of genotyping *PLA2G2A* SNP rs11573156 C>G in part of the IMPROVE study. Samples with a significant VIC signal are called homozygous for the common allele C (red cluster). A sample with an even distribution of VIC and FAM signals will be called heterozygous (green cluster). Samples with a significant FAM signal are called homozygous for the rare allele G (blue cluster). No Template Controls (NTC) are represented by black squares and should display no signal. An 'X' indicates that the sample has not reached the threshold signal level for any group

(http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_039282.pdf).

TaqMan Genotyping *PLA2G2A* SNPs of interest

All but three of 28 studies used for the Mendelian randomisation analysis (see 2.5 Table 2.6 and Table 2.7) directly genotyped rs11573156 by the genotyping methods

specifically associated with each study. Of the 3 exceptions, MedStar and Rotterdam imputed the SNP and GENDAR used a genotyped proxy. I genotyped the IMPROVE, TPT, CYPRUS, NPHSII and EAS studies using the TaqMan Genotyping Assay System and an rs11573156 specific assay (Table 2.2).

For rs3767221 all genotyping was undertaken using the TaqMan Genotyping Assay System and an rs3767221 specific assay (Table 2.2.). This SNP was genotyped previously in UDACS, EPIC-Norfolk, GRACE-France and NPHS-II. I genotyped this SNP in IMPROVE, the Edinburgh Artery Study and CYPRUS. All of these studies are detailed below in Table 2.6 and Table 2.7.

TaqMan Genotyping *PLA2G5* SNPs of Interest

TaqMan genotyping assays for the *PLA2G5* tSNPs rs11573185, rs11573191, rs11573203 and rs11573248 are defined in Table 2.2. These SNPs were genotyped in several studies (Table 2.6 and Table 2.7) as indicated in Table 2.8. Genotyping by the TaqMan method was carried out for all the non-MetaboChip studies.

Genotyping *PLA2G10* SNPs of Interest

The *PLA2G10* tSNPs rs72546339 and rs72546340 were genotyped using specific TaqMan Assays (Table 2.2). The SNPs rs4003232 and rs4003228 were genotyped by restriction fragment length polymorphism (RFLP) as described below (**2.4.6.2**). I initially attempted to genotype these SNPs using a single PCR per SNP, but this did not produce very efficient results. Therefore, nested PCR reactions were set up for each SNP by myself and Jutta Palmen.

In order to genotype rs4003228 the first reaction isolated a 2,869bp fragment and the second was a nested PCR amplification of a 220bp fragment.

PCR 1: Forward primer: 5'GCGTGACCTGCCACACCTATG3'

Reverse Primer: 5'TTCACTGGCCATGTTATCC3'

3µl of DNA (75ng) was transferred into a 96 well plate and centrifuged at 1000rpm for 1 minute to ensure that the DNA sat in the bottom of the well. Added to the DNA per reaction was: MgCl₂ 0.75µl, DMSO (5%) 1.25µl, NH₄⁺ 2.5µl, Forward primer 0.05µl, Reverse primer 0.05µl, Taq polymerase 0.1µl, dH₂O 17.3µl.

The PCR Program carried out on the G-Storm Thermocycler (Table 2.5) was: 95°C 4 minutes, then 35 cycles of; 94°C 45 seconds, 58°C 45 seconds, 68°C 3 minutes, and finally

one cycle of 72°C 5 minutes. The product was diluted with dH₂O 1:100 for use as DNA in the nested PCR. The PCR-product size was verified by electrophoresis on an agarose gel.

For rs4003228 the restriction site for Bsa AI (YAC/GTR). Below is the 220bp *PLA2G10* sequence fragment isolated by this PCR (taken from the Ensembl Genome Browser). The primers are highlighted in yellow. Blue lettering indicates exon 1 and red lettering indicates the restriction enzyme recognition site. The SNP of interest is highlighted blue. In this instance the enzyme will recognise the sequence and cut when the common allele is present but not the rare allele.

14784686	ATTTCACAGAT	GAGGAAACCAAGGCC	CAGAGAGGTTAAGAATTTGTCCAAGGTCACACAG	14784627
14784626	CCAGGAAGTAGGATTCAAACCCAGACAGCCAGGCTGTAAACACCTGGGCTCTTCTCAGGCT			14784567
14784566	CATGCCCTTCCCAGGGGTCTGGGAAGCCCTGACCTGCAGCCTGTCACCTTTGTTTACCCC			14784507
14784506	CCAG	CCTCCAGGATAT	TACGTGTCACCGGCGTGGGATCCTGGAACTGGCAGGAACTGTG	14784447

For the nested PCR 1µl of DNA was moved into a fresh 96 well plate and dried. Added to the DNA per reaction was: MgCl₂ 1.2µl, NH₄⁺ 2µl, Forward primer 0.04µl, Reverse primer 0.04µl, Taq polymerase 0.04µl, dH₂O 16.68µl. The PCR Program carried out on the G-Storm Thermocycler (Table 2.5) was: 98°C 2 minutes, then 30 cycles of; 98°C 15 seconds, 63°C 15 seconds, 68°C 20 seconds and one final cycle of; 68°C 2 minutes.

For RFLP: Per digestion reaction: Bsa AI 0.6µl, NEB Buffer 4 (10x) 1.3µl, dH₂O 8.1µl, PCR-product 3µl was added to a fresh 96 well plate. The plate was sealed and centrifuged at 1000rpm for 1 minute. The plate was then incubated at 37°C in a high humidity incubator for 16 hours. Agarose gel electrophoresis was used to check the product sizes and undigested PCR-product was used as a positive control in all instances. Expected product sizes are as follows; Homozygous common: 186+34, heterozygous: 220+186+34 and homozygous rare: 220.

To genotype rs4003232 two PCR reactions were needed The first isolated a 2,869bp fragment and the second was a nested PCR amplification of a 216bp. The first fragment isolated is the same as the 2,869bp fragment for rs4003228 (above).

For rs4003232 a recognisable restriction enzyme site for Alu I (AG/CT) was identified. The 216bp fragment of the *PLA2G10* sequence isolated by this PCR is shown below (taken from the Ensembl Genome Browser). The yellow highlight indicates the primers used. The red lettering indicates the restriction enzyme recognition site (Alu I) and the blue highlight indicates the SNP of interest. In this instance the restriction enzyme recognisable site includes the rare allele of the SNP of interest.

14784806	GTTGGGATTACAGGTGTGAGCCACTGAG	CCCAGCCGGATTATAATAGCC	TTTTTCATGCAC	14784747	
14784746	CAGGGGCTTTATACTCATTATCTCATTTTCATTCATATGAGTTGAAGTCAGTTTATCCCC			14784687	
14784686	ATTTACAGATGAGGAAACCAAGGCCAGAGAGGTTAAGAATTTGTCCAAGGTCACACAG			14784627	
14784626	CCAGGAAGTAGGATTCAAACCCAGAC	AGC(C/A)	AGGCTGTAACAC	CTGGGCTCTTCTCAGGCT	14784567
14784566	CATGC	CCTTCCCAGGGTCTGGGAAGCCCTGACCTGCAGCCTGTCACCTTTGTTTACCCC		14784507	

1µl of DNA was added to each well of a fresh 96 well plate and dried. Added to the DNA per reaction was: MgCl₂ 1.2µl, NH₄⁺ 2µl, Forward primer 0.04µl, Reverse primer 0.04µl, Taq polymerase 0.04µl, dH₂O 16.68µl. The PCR Program carried out on the G-Storm Thermocycler was: 98°C 2 minutes, then 30 cycles of; 98°C 15 seconds, 63°C 15 seconds, 68°C 20 seconds and finally one cycle of; 68°C 2 minutes.

For RFLP, per digestion reaction: Alu I 0.3µl, NEB Buffer 4 (10x) 1.3µl, dH₂O 8.4µl, PCR-product 3µl. This was incubated at 37°C in a high humidity incubator for 16 hours. Agarose gel electrophoresis was used to check the product sizes and undigested PCR-product was used as a positive control in all instances. Expected product sizes; Homozygous common: 150+66, heterozygous: 216+150+66 and homozygous rare: 216.

Micro Array Diagonal Gel Electrophoresis

Digested products were analysed by Micro Array Diagonal Gel Electrophoresis (MADGE) (7.5% gel). MADGE gels are made up in house as described in Table 2.1. The resulting arrays are analysed by eye and cross checked by two people. Genotypes were identified as follows. When the genotype is homozygous for allele 1 and the restriction site is present for both alleles it shows as 2 bands of equal intensity. This product is fully digested by the restriction enzyme. The sum of the sizes of the two bands is equal to the full product size. When the genotype is homozygous for allele 2 and the restriction site is not present for either allele the result shows as 1 band equivalent to the full product size. When the genotype is heterozygous and the restriction site is present for one allele but not the other the results shows as 3 bands. The largest band has the highest intensity and is the size of the full product. Two less intense bands that total the sizes of the fully digested product. A typical MADGE analysis of the *PLA2G10* SNP rs4003232 is shown below (Figure 2.6).

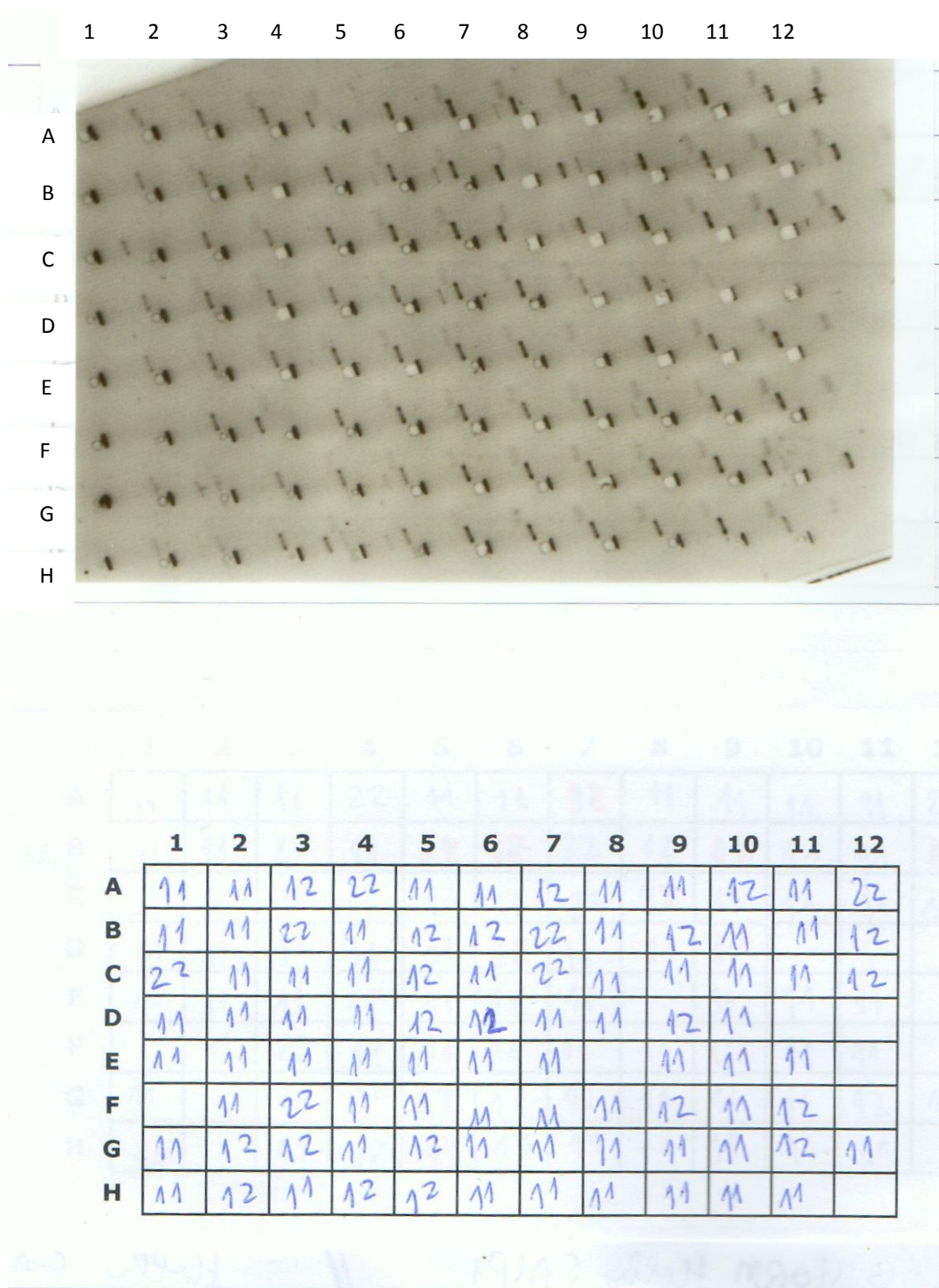


Figure 2.6. Shows restriction digested PCR samples run on a 7.5%, 96 well, MADGE gel. This gel represents the analysis of rs4003232 genotyped in samples from the EPIC-Norfolk study. The bands visible in each well recognize each sample as either: 11= homozygous for the variant with the restriction site, 22= homozygous for the variant with no restriction site or 12 as heterozygotes. The grid below the gel represents manually reported and cross checked results.

2.4.7. Luciferase Assay Analysis

The Dual-Luciferase Reporter Assay System is a genetic reporter system used to investigate eukaryotic gene expression. This assay has been used during this thesis for the purposes of comparing the luciferase expression generated by the different alleles of a SNP of interest in order to determine if there is a differential genotype effect. This can be achieved by creating DNA fragments containing each allele of the SNP of interest and inserting them into a pGL3-Luciferase Reporter Vector. The 'dual' system refers to the concurrent expression and measurement of two individual reporter enzymes in the same system. There was an experimental reporter enzyme and a co-transfected control reporter enzyme. The experimental enzyme was expressed based on the specific experimental conditions, whereas the control enzyme activity was measured as the internal baseline response and was used to normalize the experimental reporter activity, thus increasing the efficiency of the assay. There were several preparatory stages that must be completed before a luciferase assay is run.

2.4.7.1. Isolating a DNA Fragment for Cloning into a pGL3-Luciferase Reporter Vector

The vector and fragment were accomplished by digestion of the PCR-product and the vector at the polylinker site with the appropriate restriction enzymes to synthesize compatible 'sticky' ends for cloning. These reactions were carried out as per restriction enzyme manufacturers' instructions. Completion of each double digestion was verified by agarose gel electrophoresis. Undigested sample of the original PCR-product and undigested pGL3-Luciferase Reporter Vector DNA were used as controls. If the digestion was successful a single clean band that was demonstrably smaller than the original PCR-product would be visible when the gel was analysed

(<http://www.promega.com/~media/Files/Resources/Protocols/Technical%20Manuals/0/pGL3%20Luciferase%20Reporter%20Vectors%20Protocol.ashx>).

2.4.7.2. Ligation of Digested Vector and DNA Fragment (Insert)

To ligate the compatible 'sticky' ends of the vector polylinker site and insert, the T4 DNA ligase kit was used. Each 20µl total reaction mix was set up and incubated on ice. Each component was added in the following specific order:

50ng (0.025pmol) Vector DNA (3kb), then 50ng (0.076pmol) Insert DNA (1kb), then 2µl T4 DNA Ligase Buffer (10x) (thawed and re-suspended at room temperature), then up to 19µl dH₂O was added and finally 1µl T4 DNA Ligase.

The reaction mix was gently pipetted up and down, briefly centrifuged and incubated at 16°C overnight (<http://www.neb.com/nebecomm/products/protocol658.asp>).

2.4.7.3. Transforming Vector/Insert Complex into Competent Cells

For rapid amplification of the vector/insert complex it was transformed into *E.coli* DH5α cells as follows: 50µl DH5α cells thawed on ice and added to 1-5µl ligated DNA (1-10ng). This was mixed by gently pipetting up and down and incubated on ice for 30 minutes, while pre-warming 300µl LB Medium to 37°C. Cells were heat shocked for 20 seconds in a 42°C water-bath. On removal from the water-bath the 300µl of pre-warmed LB was immediately added to the cell/DNA mix. The mix was then incubated on a shaker at 37°C/225rpm for 1 hour and spread onto a fresh plate of LB agar + ampicillin (50mg/ml) (Table 2.1). Finally the plate was incubated at 37°C for 16 hours.

2.4.7.4. Amplification of Transformed Colonies and DNA Purification by Miniprep

A single colony was picked from each agarose plate and amplified in 5ml sterile LB medium and inoculated with 10µl of the antibiotic ampicillin (50mg/ml). The pGL3-Luciferase Reporter Vector (Figure 2.3) contains an ampicillin resistance gene. The culture was incubated on a shaker for 16 hours at 37°C/225rpm. Following amplification of the colony, glycerol stocks of the culture were made: 1ml of the culture was added to 1ml of glycerol (50%) and stored at -80°C for future experiments. Plasmid DNA was purified using the QIAprep Spin Miniprep Kit purified up to 20µg of high-copy plasmid DNA from 1-5ml of *E. coli* culture. The bench protocol for the QIAprep Spin Miniprep Kit using a microcentrifuge as per manufacturers' instructions. (www.qiagen.com/hb/qiaprepminiprep).

Plasmid Design for rs11573156 C>G

Previous studies have suggested that regulatory elements for *PLA2G2A* start occurring approximately 2.27Kb upstream of the translation start site (Cannon 1998), this is around 600bp upstream from the start of the suggested 5'UTR. Considering this, a 2.27Kb fragment upstream of the *PLA2G2A* translation start site (GRCh37.p5, chr1:20,305,406 to 20,307,676) (Ensembl Genome Browser), including wild-type rs11573156 C, was amplified from a previously genotyped DNA sample of an individual homozygous for the C allele. The DNA used was from an in-house cohort known as the DRD study, a small, volunteer cohort of DNA taken from individuals working at the University College London Rayne Building or living in close proximity. The study is composed of males and females between the ages of 19 and 65 years.

The primers used to create DNA fragments to insert into a plasmid for use in luciferase assays included two restriction endonuclease recognition sites (one forward and one reverse) to pair with the polylinker sites in the pGL3 vector (Figure 2.1). The primers included restriction endonuclease sites for NheI and HindIII and were as follows:

Forward, 5' CCCC GCTAGCTGATCTCTGCCTTCATCTTTGTATG

Reverse, 5'CCCC AAGCTTCTGCTGGGTGGTCTCAACTTC.

The PCR conditions were; 95°C for 1 minute, then 30 cycles of 95°C for 1 minute, 68°C for 1 minute and 72°C for 1 minute, ending with 72°C for 5 minutes. The amplified fragment was ligated into the pGL3-basic (Figure 2.1) vector upstream of the Luciferase (*Luc*) gene, using the HindIII and NheI polylinker sites and transformed into *E.coli* DH5α competent cells.

Plasmid Design for rs3767221 A>C

A 303bp fragment downstream of the *PLA2G2A* protein coding sequence (GRCh37.p5, chr1:20,301,598 to 20,301,873) (Ensembl Genome Browser) was identified to include rs3767221. As shown in Figure 1.18 the 3'UTR of *PLA2G2A* is ~263bp, therefore sequence of around 300bp was isolated since this would include the whole 3'UTR for *PLA2G2A*. The sequence was amplified using the primers, which included the restriction endonuclease sequences for BamHI and Sall:

Forward 5'CGCGGATCCCATGCAGGAGGCACCAAGTGTATCTC

Reverse 5'ACGCGTCGACGCATGAGAACAACAGGGGTAGGGG

Amplification of DNA was from two previously genotyped individual DRD samples; a sample homozygous for the wild-type allele (A) and a sample homozygous for the rare variant (C). The PCR conditions were; 98°C for 2 minutes, then 30 cycles of 98°C for 15 seconds, 62°C for 15 seconds and 72°C for 1 minute, ending with 72°C for 5 minutes.

PCR fragments were ligated into a pGL3-Promoter expression vector (Figure 2.2), downstream of the *Luc* gene, using the BamHI and Sall polylinker sites. Constructs were transformed into DH5α competent *E.coli*. All products were sequenced for verification.

2.4.7.5. Verification of Amplification of the Correct DNA Sequence

To verify that the correct plasmid DNA had been amplified, a portion of the miniprep product was digested with the appropriate restriction enzymes, as per the restriction enzyme manufacturers' instruction.

To ensure that the sequence of the DNA was identical to the consensus sequence a sample of the miniprep solution was sent for sequencing. In 3 separate eppendorf tubes, 10µl DNA (~100ng/µl) and two insert-flanking primers (33.3pmol per primer) were sent to Source BioScience UK. Sequencing results were sent as both a text file and an electropherogram, visible using FinchTV (Table 2.4). Where the software was unable to identify a base, the peaks of the sequence were checked by eye to identify whether there was a true error or not. The text file was then used to compare the sequenced DNA to the consensus sequence (Ensembl Genome Browser) with the online tool CLUSTALW to highlight discrepancies. If a discrepancy is found to be a true mistake it was rectified using the Quikchange site-directed mutagenesis kit.

To correct the sequence the kit used the supercoiled double-stranded DNA vector/insert complex and two synthetic oligonucleotide primers (one forward and one reverse) with the desired mutation. The DNA polymerase worked to extend the primers during a PCR and their incorporation synthesized mutated plasmids containing staggered nicks. Post-PCR the product was treated with *Dpn*I endonuclease, which is specific for methylated and hemimethylated DNA and digested the parental DNA template, selecting for the newly formed, corrected DNA sequence. This DNA was then transformed into XL1-Blue supercompetent cells, amplified and purified by miniprep as previously described (<https://www.life-science-lab.org/.../ag-biochem-projekt2006>).

Site directed mutagenesis was required to correct mistakes in the original, wild type rs11573156 (suspected) promoter fragment (which was used in luciferase assays) as shown below. The CLUSTAW tool (Table 2.4) was used to compare the Ensembl consensus sequence with the sequenced clone fragment text file (provided by Geneservice (Table 2.4)). A match is indicated by a star below each comparable allele. Where a gap occurs there was a mismatch in the sequence. A mismatch may represent a known polymorphism or a mistake. In the sequence below mistakes changed by site directed mutagenesis are highlighted in yellow and the corresponding primers are listed below the fragment.

Clone PLA2G2A promoter sequence: All mistakes corrected by SDM.

```
>Ensemble<<      TGATCTCTGCCTTCATCTTTGTATGGCCTTCTCCAGTGTGCCTGCCTGTCTCCAAATTT
>Clone_PLA2G2A_full_seq<< TGATCTCTGCCTTCATCTTTGTATGGCCTTCTCCAGTGTGCCTGCCTGTCTCCAAATTT
*****

>Ensemble<<      CCCATTTTGTAAATGACATCTGTCTATCTTGGATTAGGGCCCATCCTAGCTGCCCTAGTGA
>Clone_PLA2G2A_full_seq<< CCCATTTTGTAAATGACATCTGTCTATCTTGGATTAGGGCCCATCCTAGCTGCCCTAGTGA
*****

>Ensemble<<      TCTCCTCTTAACTAATTATATCTGCAATGACCCCTATTTTAAATAAGATCACATTCTAAG
>Clone_PLA2G2A_full_seq<< TCTCCTCTTAACTAATTATATCTGCAACGACCCCTATTTTAAATAAGATCACATTCTAAG
*****

>Ensemble<<      GTACCGGGACTCCAACACAGGAAGTGTAGAGCAGGAAGGAGGACCACAGTCCAACCCCTTA
>Clone_PLA2G2A_full_seq<< GTACCGGGACTCCAACACAGGAAGTGTAGAGCAGGAAGGAGGACCACAGTCCAACCCCTTA
*****

>Ensemble<<      CACAGTCGAGCACTGGAGGTAATGGTATTCTCGAGAGGAAAGCTGTTTGATGGCTCTTG
>Clone_PLA2G2A_full_seq<< CACAGTCGAGCACTGGAGGTAATGGTATTCTCGAGAGGAAAGCTGTTTGATGGCTCTTG
*****

>Ensemble<<      GGATAGATTAGGTTATGCTCCAGCAACTGCCCATCTCAATTATCATCAGCTTAATACAAC
>Clone_PLA2G2A_full_seq<< GGATAGATTAGGTTATGCTCCAGCAACTGCCCATCTCAATTATCATCAGCTTAATACAAC
*****

>Ensemble<<      AAAAGTCTATTTCTCACTCACACTAAGACCCCTGTGAGCTGCAGGACTTTCTAGGACAAC
>Clone_PLA2G2A_full_seq<< AAAAGTCTATTTCTCACTCACACTAAGACCCCTGTGAGCTGCAGGACTTTCTAGGACAAC
*****

>Ensemble<<      TGTCTTCCGTGCCATGATTCTTACACCCAGATTGCTTTCATCTTCTGGTTCTACCAAGTG
>Clone_PLA2G2A_full_seq<< TGTCTTCCGTGCCATGATTCTTACACCCAGATTGCTTTCATCTTCTGGTTCTACCAAGTG
*****

>Ensemble<<      GGGCAGGGAAGGGCTGGGGAGTTGCATCTCATAGTCCCTTGCTGGAGAAACCTGGAAGT
>Clone_PLA2G2A_full_seq<< GGGCAGGGAAGGGCTGGGGAGTTGCATCTCATAGTCCCTTGCTGGAGAAACCTGGAAGT
*****

>Ensemble<<      GACCCCTTCATACCAGTGTCTATGTTTACCAGTCAGCAGCAGGGCCCTAGGTCTCCATCC
>Clone_PLA2G2A_full_seq<< GACCCCTTCATACCAGTGTCTATGTTTACCAGTCAGCAGCAGGGCCCTAGGTCTCCATCC
*****

>Ensemble<<      CTGCAAGGGGCTCTAGAGTGTAGGGAACAGATGGAATGTTTACAAGCACCACAGCCCTG
>Clone_PLA2G2A_full_seq<< CTGCAAGGGGCTCTAGAGTGTAGGGAACAGATGGAATGTTTACAAGCACCACAGCCCTG
*****

>Ensemble<<      CCACAGTGACTTTTTTAGGACTGGTATCGCAGAGTGTTTACTTTAAGGCGGTGGAAGCTAA
>Clone_PLA2G2A_full_seq<< CCACAGTGACTTTTTTAGGACTGGTATCGCAGAGTGTTTACTTTAAGGCGGTGGAAGCTAA
*****
Position 780 corrected g>a
ATTCTTAGCATGTGCTGGAGAGCAATGAAAAGATATTTACTTTATGAATTAAAGCTGGA
>Ensemble<<      ATTCTTAGCATGTGCTGGAGAGCAATGAAAAGATATTTACTTTATGAATTAAAGCTGGA
>Clone_PLA2G2A_full_seq<< ATTCTTAGCATGTGCTGGAGAGCAATGAAAAGATATTTACTTTATGAATTAAAGCTGGA
*****

>Ensemble<<      GTCAGTGTGAGCCCGAAGGCTGAAGGAAAAAGAGCAACAGATCCAGGGAGCATTACCTG
>Clone_PLA2G2A_full_seq<< GTCAGTGTGAGCCCGAAGGCTGAAGGAAAAAGAGCAACAGATCCAGGGAGCATTACCTG
*****

>Ensemble<<      CCCTGTCTCCAAACAGGTGAGGATGGGGAATAAAGTGAAGGCAGTGTCTTGGTGGGAAGT
>Clone_PLA2G2A_full_seq<< CCCTGTCTCCAAACAGGTGAGGATGGGGAATAAAGTGAAGGCAGTGTCTTGGTGGGAAGT
*****

>Ensemble<<      TCAAGGATAGCCTCTGGCTTTTCCAGGTTTAGAAGCTCATATGAGACAGGGGTGGAGGA
>Clone_PLA2G2A_full_seq<< TCAAGGATAGCCTCTGGCTTTTCCAGGTTTAGAAGCTCATATGAGACAGGGGTGGAGGA
*****

>Ensemble<<      AAAGAAGAAAGAAGAATAAGAAGAGAAAGTTGAGGCCCTGGCCCAAGTTAGTGGGAAGGG
>Clone_PLA2G2A_full_seq<< AAAGAAGAAAGAAGAATAAGAAGAGAAAGTTGAGGCCCTGGCCCAAGTTAGTGGGAAGGG
*****

>Ensemble<<      AAATCCACCCCATTAACCTCTCTCCCTGGTGGACTGTGGGGTGCACACGTGAGGCCTGC
>Clone_PLA2G2A_full_seq<< AAATCCACCCCATTAACCTCTCTCCCTGGTGGACTGTGGGGTGCACACGTGAGGCCTGC
*****

>Ensemble<<      ACAGGTGCTGGAACATGGTAGAGGCCAGGACATACTTCCTGTGAATGAATGATTGAGCG
>Clone_PLA2G2A_full_seq<< ACAGGTGCTGGAACATGGTAGAGGCCAGGACATACTTCCTGTGAATGAATGATTGAGCG
*****

>Ensemble<<      GCTGAATGAATGAGTACCACATAAAGCCCTCTTTCTATTCCCAAATGCCACATTGAGCA
>Clone_PLA2G2A_full_seq<< GCTGAATGAATGAGTACCACATAAAGCCCTCTTTCTATTCCCAAATGCCACATTGAGCA
*****

>Ensemble<<      GAAGGGAGCAATCCTTGCTCAGCAATTGGTAGTCCCTTTGGGTGTGCAAATGAGTCCACA
>Clone_PLA2G2A_full_seq<< GAAGGGAGCAATCCTTGCTCAGCAATTGGTAGTCCCTTTGGGTGTGCAAATGAGTCCACA
*****

>Ensemble<<      GCCTGCAACAGCAGACAGTCTCTGCCCCCTTAGAGGCGATTGCAGGGAGGTGGCTGACC
>Clone_PLA2G2A_full_seq<< GCCTGCAACAGCAGACAGTCTCTGCCCCCTTAGAGGCGATTGCAGGGAGGTGGCTGACC
*****
```

```

>Ensemble<<                               Position 1369 corrected c>t
>Clone_PLA2G2A_full_seq<<                GTTGTATCACACCCAGAGCTGGTTATGGGAATTTACTCCATGGAAAGACTGCAAAACTGCC
*****
>Ensemble<<                               TGAATGTGTTTTTGGCATCAGCTACTGACACGTAAGGTTTCCCAATCCTCAACTCTGTCC
>Clone_PLA2G2A_full_seq<<                TGAATGTGTTTTTGGCATCAGCTACTGACACGTAAGGTTTCCCAATCCTCAACTCTGTCC
*****
>Ensemble<<                               TGCCAGCTGATGAGGGGAAGGAAAGGGATTACCTAGGGGTATGGGCGACCAATCCTGAGT
>Clone_PLA2G2A_full_seq<<                TGCCAGCTGATGAGGGGAAGGAAAGGGATTACCTAGGGGTATGGGCGACCAATCCTGAGT
*****
>Ensemble<<                               CCACCAACTGACCACGCCCATCCCCAGCCTTGTGCCTCACCTACCCCCAACCTCCCAGAG
>Clone_PLA2G2A_full_seq<<                CCACCAACTGACCACGCCCATCCCCAGCCTTGTGCCTCACCTACCCCCAACCTCCCAGAG
*****
>Ensemble<<                               GGAGCAGCTATTTAAGGGGAGCAGGAGTGCAGAACAAACAAGACGGCCTGGGGATACAAAC
>Clone_PLA2G2A_full_seq<<                GGAGCAGCTATTTAAGGGGAGCAGGAGTGCAGAACAAACAAGACGGCCTGGGGATACAAAC
*****
>Ensemble<<                               TCTGGAGTCCTCTGAGAGGTAAGAGAGCCAGCGAAGCTGATGTCCTGTCAAGAGCAGAATT
>Clone_PLA2G2A_full_seq<<                TCTGGAGTCCTCTGAGAGGTAAGAGAGCCAGCGAAGCTGATGTCCTGTCAAGAGCAGAATT
*****
>Ensemble<<                               CCTGCTCATTGCTGCCTTTGAGAGTGGCTGTGTGTGCATGCATGTGCATGATTGATA
>Clone_PLA2G2A_full_seq<<                CCTGCTCATTGCTGCCTTTGAGAGTGGCTGTGTGTGCATGCATGTGCATGATTGATA
*****
>Ensemble<<                               TGTATGAGAGGGTGTGTGTGCATGAGTGTGTGAGTGAGTATGTGAGTGTAGTGTAAAGAG
>Clone_PLA2G2A_full_seq<<                TGTATGAGAGGGTGTGTGTGCATGAGTGTGTGAGTGAGTATGTGAGTGTAGTGTAAAGAG
*****
>Ensemble<<                               Known polymorphism
>Clone_PLA2G2A_full_seq<<                AGGATGTTGGCACTATCAGGTAACTACGAGAGTGTGTGTATGTGGGCATAGGTGTGTTAA
AGGATGTTGGCACTATCAGGTAACTACGAGAGTGTGTGTATGTGGGCATAGGTGTGTTAA
*****
>Ensemble<<                               CATGTATGTGTTTGGGAACTTGTGTATGTGGAAGGGGTTAGAAGGCCTAGAAGAGAGAGG
>Clone_PLA2G2A_full_seq<<                CATGTATGTGTTTGGGAACTTGTGTATGTGGAAGGGGTTAGAAGGCCTAGAAGAGAGAGG
*****
>Ensemble<<                               TTGATGCTTTTCATTCTGGAGGAAAATACTGAGGCCGAGCCTCCATGGGTGCCTTGGAGAC
>Clone_PLA2G2A_full_seq<<                TTGATGCTTTTCATTCTGGAGGAAAATACTGAGGCCGAGCCTCCATGGGTGCCTTGGAGAC
*****
>Ensemble<<                               TCCAAGCCTTGAATCCAGTGTGGGGATATGCAAGCTATGTCTAGCGAGGGACACATCCTC
>Clone_PLA2G2A_full_seq<<                TCCAAGCCTTGAATCCAGTGTGGGGATATGCAAGCTATGTCTAGCGAGGGACACATCCTC
*****
>Ensemble<<                               TGACCTCAGGAACCTCCCAGGTAGTTGGGAGGAACCTGGTTCCAACCTCCCAAGAACTCT
>Clone_PLA2G2A_full_seq<<                TGACCTCAGGAACCTCCCAGGTAGTTGGGAGGAACCTGGTTCCAACCTCCCAAGAACTCT
*****
>Ensemble<<                               position 2125 corrected c>t
>Clone_PLA2G2A_full_seq<<                CAGTCTGATGAGGTACAGGGGAGGTCTCATTAGTGTATCATGGGGTTCTCCACAGGTCTG
CAGTCTGATGAGGTACAGGGGAGGTCTCATTAGTGTATCATGGGGTTCTCCACAGGTCTG
*****
>Ensemble<<                               AGGGCCTGATGTGTGTGAAACCATTTCTGCAGAGCTGGGAACGGGTGAGGAGGTGGTTGTG
>Clone_PLA2G2A_full_seq<<                AGGGCCTGATGTGTGTGAAACCATTTCTGCAGAGCTGGGAACGGGTGAGGAGGTGGTTGTG
*****
>Ensemble<<                               Known repetition
>Clone_PLA2G2A_full_seq<<                TGTGTGTGTGTGTGTGTGTGTGTGTCATTGCTGGAGGGCACTCCTTGTGTGCTCTGAGTGT
TGTGTGTGTGTGTGTGTGTGTGTGTGTCATTGCTGGAGGGCACTCCTTGTGTGCTCTGAGTGT
*****
>Ensemble<<                               Position 2300 corrected insert g
>Clone_PLA2G2A_full_seq<<                GACAGAGGAAGTCACCTGGACTTAGGTTGGATGGGAGAGCATGTCGTGTGTCTCAGAG
GACAGAGGAAGTCACCTGGACTTAGGTTGGATGGGAGAGCATGTCGTGTGTCTCAGAG
*****
>Ensemble<<                               CCACCAAGGAGGAGCAGGGGAGCGACGGCCGGGGCAGAAGTTGAGACCACCCAGCAG
>Clone_PLA2G2A_full_seq<<                CCACCAAGGAGGAGCAGGGGAGCGACGGCCGGGGCAGAAGTTGAGACCACCCAGCAG
*****

```

Primers for these site directed mutagenesis changes are shown below written stating the incorrect allele, the position of the allele and the corrected allele:

G780A Forward: 5' TTTATGAATTAAGCTGGAGTGTGACCCCGAAG 3'
G780A Reverse: 5' CTTGGGCTGACACTGACTCCAGCTTTAATTCATAAA 3'

C1369T Forward: 5'AATTTACTCCATGGAAAGACTGCAAACTGCCTGAAATGTG '3
C1369T Reverse: 5' CACATTTTCAGGCAGTTTTGCAGTCTTCCATGGAGTAAATT '3

C2125T Forward: 5' TGAGGTACAGGGGAGGTCTCATTAGTGTATCATG '3
C2125T Reverse: 5' CATGATACACTAATGAGACCTCCCCTGTACCTCA '3

Ins2300G Forward: 5' AGAGGAAGTCACCCTGGACTTAGGTTGGATGGG '3
Ins2300G Reverse: 5' CCCATCCAACCTAAGTCCAGGGTGAATTCCTCT '3

PLA2G2A rs11573156 Site Directed Mutagenesis to Introduce a Base

Site directed mutagenesis was also used to deliberately introduce the rs11573156 G allele to the plasmid DNA insert sequence, for comparison to the wild-type sequence, under the following conditions:

Forward primer: 5'CCTCCCAGAGGGAGGAGCTATTTAAGGGG

Reverse primer: 5'CCCCTTAAATAGCTCCTCCCTCTGGGAGG

Primers were used at 125ng/μl. Each reaction included; 5μl Reaction Buffer (10x), 1μl dNTPs, 1.5μl QuikSolution Reagent, 1μl Forward primer, 1μl Reverse primer, 1μl DNA (50ng/μl), 39.5μl dH₂O, then add 1μl QuikChange Lightning Enzyme. The PCR conditions were as follows: 95°C for 2 minutes, then 18 cycles of; 95°C for 20 seconds, 60°C for 10 seconds and 68°C for 4.5 minutes (30 seconds/Kb plasmid length) followed by a final cycle of; 68°C for 5 minutes. The PCR-product was then digested with 2μl DpnI and underwent 10 minutes of incubation at 37°C. The DpnI treated product was then re-transformed into XL1-Blue supercompetent cells.

2.4.7.6. Amplification of Re-grown Colonies and DNA Purification by Maxiprep

I used DNA at a concentration of 400ng/μl for luciferase assays. In order to achieve this concentration, more DNA was amplified and purified. The purification process was by maxiprep. This was achieved as follows: A glycerol stock of the cell culture was thawed on ice and a sample of cells was streaked onto a fresh LB agar plate (Table 2.1). The remaining cells were re-stored at -80°C.

The plate was incubated at 37°C for 16 hours after which a fresh colony was picked and added to 150ml of sterile LB medium and inoculated with 300μl of ampicillin (100μg/ml). This was incubated on a shaker for a further 16hours at 37°C/225rpm. Maxiprep reactions were carried out as according to the manufacturers' guidelines.

<http://www.sigmaaldrich.com/etc/medialib/docs/Sigma/Bulletin/na0300bul.Par.0001.File.tmp/na0300bul.pdf>.

2.4.7.7. Transfection of Purified Plasmid into a Human Cell Line.

Luciferase assay analysis was carried in order to measure the luciferase activity generated by the plasmid/insert complex. Human hepatocellular carcinoma cell line (Huh7 cells) (Table 2.1) were incubated at 37°C in DMEM (+10% FBS) (Table 2.1) and grown to 95% confluency in a t-75 flask. Trypsin (0.25%) (Table 2.1), PBS (1x) and fresh DMEM were then pre-warmed to 37°C. The Huh7 cells were washed with 10ml PBS (1x) after which the PBS and DMEM were drained. The cells were trypsinised with 3ml trypsin for 2 minutes at 37°C to remove cells from the flask wall. They were then immediately washed with 12ml fresh DMEM and transferred to a 15ml falcon tube. The cells were then centrifuged at 1,000g for 5 minutes and the supernatant was discarded. The cell pellet was resuspended in 12ml DMEM. The cell count was estimated using the Advanced Detection and Accurate Measurement (ADAM) cell counter.

Cells were plated in a 96 well format. 100µl of the cell solution was added to each well (2×10^4 cells/100µl) and incubated at 37°C for 24 hours. For each luciferase assay carried out in this thesis the plate would include one column containing the puc 19 plasmid, one containing pGL3-Basic plasmid, one containing pGL3-control plasmid and one containing β-Gal. Puc 19 is a plasmid that does not contain the luciferase (*luc+*) gene. pGL3-Basic is a plasmid that contains both the *Renilla* and Firefly *luc+* gene, but no promoter sequence. pGL3-Basic luciferase measurements were used to normalise all vector/insert plasmid measurements, so different DNA plasmids could be compared effectively within and between each plate. pGL3-Control is a plasmid that contains the *luc+* gene and an internal promoter (the SV40 promoter) sequence, known to drive luciferase activity to high levels. This was used as a positive control for luciferase activity. β-Gal is an enzyme that catalyses the hydrolysis of β-galactosidase into monosaccharides. The β-Gal assay was used to detect transfection efficiency. Detection was by X-gal when X-gal was cleaved by active β-Gal enzymes in this reaction it synthesized a blue product, easily identified by eye or under a light microscope.

For 1x 96 well plate, cells were transfected for luciferase assays as follows: A mastermix of 12µl of pRL-TK (10ng/µl) and 2,928µl of Opti-MEM was made and mixed by inverting 8-10 times. PRL-TK activated the *renilla* luciferase gene in the plasmid DNA. This measure was the co-transfected control reporter enzyme activity.

5µl of plasmid DNA (400ng/µl) and 245µl of the mastermix was added to an eppendorf tube, except in the case of puc plasmid DNA and β-Gal enzyme, which do not require the pRL-TK co-transfector as they have no *luc+* genes. In this case 245µl of Opti-MEM was added to each eppendorf tube.

A lipid mastermix of 66µl of Lipofectamine 2000 (1mg/ml) was added to 3,243µl of Opti-MEM and mixed by inverting 8-10 times. The mix was then left to stand for 5 minutes. Lipofectamine 2000 is a transfection reagent that contains lipid subunits capable of forming liposomes in an aqueous environment, which trap the DNA plasmids and then fuse to the plasma membrane of the Huh7 cells and are engulfed. 250µl of lipid mix was added to each eppendorf tube to complete the transfection mix. Each transfection mix was left at room temperature for 20 minutes. 50µl of transfection mix was added to each well of the 96 well plate and the plate was incubated at 37°C for 48 hours.

2.4.7.8. Cell Lysis and β -Gal Assay

For the β -Gal assay, the medium was removed from the cells and washed with 100µl PBS (1x). 50µl of Fixture Solution per well was added and left at room temperature for 5 minutes. The Fixture Solution was removed and the cells were washed twice with 100µl of PBS (1x). 50µl of Staining Solution was added to each well.

To lyse the cells in preparation for the luciferase assay the medium was removed from all wells and the cells were washed twice with PBS (1x). Then 20µl of Lysis Buffer (1x) was added to all wells and the plate was gently rocked for 15 minutes at room temperature.

2.4.7.9. Luciferase Assay

The Dual-Luciferase Reporter Assay System sequentially measured the activity of both firefly and *renilla* luciferases from a single sample. The firefly was measured by the addition of Luciferase Assay Reagent II (LAR II) to generate a stabilized luminescent signal. After quantification of this luminescent signal this reaction was quenched and simultaneously the *renilla* reaction was initiated by the addition of the Stop & Glo Reagent to the same well. Stop & Glo also produced a stabilized signal from the *renilla* luciferase, which decayed over the course of the measurement.

The Dual-Luciferase Reporter Assays (Table 2.1) were run using the Tropix TR717 Microplate Luminometer (Table 2.5) using WinGlow software (Table 2.4). The machine injectors were set to dispense 100µl of LAR II and Stop & Glo respectively. All measurements were set for a 2 second delay and a 10 second read time. 10µl of lysed cells from each well were transferred by pipette to a white cell culture plate and placed into the Luminometer. LAR II was dispensed into the first well and the firefly luciferase activity was measured and recorded. Stop & Glo was dispensed and the *renilla* activity was measured and recorded. This process was repeated for each well.

The resulting measurements were recorded using the WinGlow software (Table 2.4). The ratio of Firefly:*renilla* activities was used to effectively compare different samples and plates.

2.4.8. LightShift Chemiluminescent Electrophoretic Mobility Shift Assay (EMSA)

The purpose of an electrophoretic mobility shift assay (EMSA) is to identify any specific protein complexes binding to a sequence of interest. Comparison of EMSA results for the genotypes of a SNP of interest can identify potential differential transcription factor, which may have a regulatory effect on the transcription of the gene of interest.

Throughout this thesis the chemiluminescent reaction is used, as opposed to the radiolabelled EMSA reaction. The chemiluminescent reaction works by detecting biotin labelled primers using the stabilized Streptavidin-Horseradish Peroxidase detection.

As a control in each experiment, each set of biotin labelled primers was run alongside a reaction that had a 20x the same, unlabelled primers added to the reaction before the addition of the biotin labelled primers. As the unlabelled primers were added before the labelled primers, they annealed themselves to any specific protein-DNA binding sites, preventing the labelled primers from annealing there, thus 'competing' out the labelled primers. Non-specific binding would not be competed out by the specific, unlabelled primers.

2.4.8.1. Designing, Biotin Labelling and Annealing Primers

EMSA Probes for the Analysis of rs11573156

Forward and reverse primers were designed, as previously described, to be ~30 base pairs in length. EMSA probes used for rs11573156 C>G were as follows:

Forward wild type 5'-AACCTCCCAGAGGGAGCAGCTATTTAAGGGGAG-3'

Reverse wild type 5'-CTCCCCTTAAATAGCTGCTCCCTCTGGGAGGTT-3'

Forward rare variant 5'-AACCTCCCAGAGGGAGGAGCTATTTAAGGGGAG-3'

Reverse rare variant 5'-CTCCCCTTAAATAGCTCCTCCCTCTGGGAGGTT-3'

EMSA Probes for the Analysis of rs3767221

EMSA probes used for rs3767221 A>C were as follows:

Forward wild type 5'-TGAGCTCAAGCAATCATTGCACTTCAGCCT-3'

Reverse wild type 5'-AGGCTGAAGTGCAAIGATTGCTTGAGCTCA-3'

Forward rare 5'-TGAGCTCAAGCAATCCTTGCACTTCAGCCT -3'

Reverse rare 5'-AGGCTGAAGTGCAAGGATTGCTTGAGCTCA -3'

Each primer was diluted to 5µM for biotin labelling. For a standard 50µl biotin labelling reaction: 29.5µl dH₂O, 10µl TdT Buffer (5x) (Table 2.1), 5µl biotin-11-dUTP (Table 2.1), 0.5µl TdT (Table 2.1) and 5µl primer (5µM).

This mix was incubated at 37°C for 90 minutes and then at 70°C for 10 minutes to stop the reaction. An equal volume of chloroform:isoamyl alcohol (24:1) was added to extract TdT (50µl) and this mix was vortexed briefly and centrifuged at 13,000rpm for 2 minutes. The top (aqueous) layer was retained.

To anneal both the biotin labelled and unlabelled primer sets (forward and reverse) equal volumes of labelled forward and reverse primers (5µM), or unlabelled forward and reverse primers (100pMol/µl) were mixed together. 1/10th the total volume of annealing buffer (10x) was added and vortexed briefly. Primers then underwent an annealing reaction on the G-Storm thermocycler: 95°C for 5 minutes, 70-95°C (-1°C/cycle) for 1 minute and then 4°C HOLD. Primers were stored at -80°C in 50µl aliquots.

2.4.8.2. Nuclear Extraction

Mammalian eukaryotic cell nuclear extract contains protein-DNA complexes and were isolated from Huh7 cells grown to 95% confluency in a t-175 flask as follows: Cells were washed with 25ml PBS (1x) and trypsinised with 5ml trypsin (0.25%) at 37°C for 3 minutes. 30ml DMEM was added to the cells and the mix was transferred to a falcon tube and centrifuged at 4°C for 5 minutes at 1000rpm. The supernatant was discarded. The cell pellet was resuspended in 5ml ice-cold Buffer A with 50µl of protease inhibitor (100x). This mix was incubated on ice for 10 minutes. The mix was then centrifuged at 4°C for 5 minutes at 1000rpm. The supernatant was discarded. The cell pellet was resuspended in 2ml of ice-cold Buffer A with 20µl of protease inhibitor (100x). The mix was then vortexed for 30 seconds. This mix was centrifuged at 4°C for 2 minutes at 13,000rpm. The supernatant was discarded. The cell pellet was resuspended in 800µl of ice-cold Buffer C with 16µl of protease inhibitor (100x). The mix was vortexed for 1 minute and incubated on ice for 10 minutes. This vortex/incubation process was repeated 4 times in total. The mix was then centrifuged at 4°C for 50 minutes at 13,000rpm. The supernatant was aliquoted into 50µl volumes and stored at -80°C.

2.4.8.3. Polyacrylamide Gel (6%) Preparation

Chemiluminescent EMSA reactions were run by gel electrophoresis on a 6% polyacrylamide gel (1x1.5mm). A frame was constructed to hold two clean glass panels with

a 1.5mm gap for the gel to be poured into. The gel was poured immediately into the gap between the glass panels and the dividing comb was placed into the top of the gel between the glass panels to create the well spaces. The gel was left to polymerise.

Once polymerised, the comb was removed from the gel and the wells were washed with TBE (0.5x). Then 1µl of EMSA Loading Buffer was added into each well and the gel was moved to a cold room (4°C). The gel was placed in a tank containing TBE (0.5x) and electrophoresed at 120V for 1 hour to ensure that everything was working.

2.4.8.4. EMSA Binding Reaction and Gel Electrophoresis

There were two EMSA binding reactions, one for the competitor reactions and one for the non-competitor reactions. For both reactions: 2µl Binding buffer (10x) (Table 2.1), 1µl poly dIdC (Table 2.1), 0.5µl MgCl₂ (50mM) (Table 2.1), 1.5µl Random primers (Table 2.1), 3µl nuclear extract. Then for the competitor reactions the following was added: 6µl dH₂O, 4µl unlabelled primers. The competitor reactions were incubated at 4°C for 30 minutes. Then for non-competitor reactions 10µl dH₂O was added and for both reactions 2µl biotin labelled primers was added and all reactions were then incubated at 25° for 50 minutes. 4µl of EMSA loading buffer was then added to each reaction and before loading into the pre-prepared gel and run at 120V for 5 hours.

2.4.8.5. EMSA Blotting Membrane

Once gel electrophoresis was complete the protein was blotted onto a nylon membrane for the purposes of visualization. 4x pieces of Whatman paper were cut to ~15cm x 10cm, then a piece of Hybond-N+ membrane was also cut to the same dimensions. These were all soaked in SSC (10x) for 10 minutes.

A tank containing SSC (20x) with a Perspex block in the middle was assembled. Over this block 2 pieces of Whatman paper were layered so that the ends were submerged in the SSC. This served as the surface to layer the blotting membrane assembly onto and kept all the layers moist. Assembly included: 2 pieces of SSC soaked Whatman paper which the gel was laid on top of. The edges of the gel were then covered with Clingfilm. The Hybond-N+ membrane was layered next, ensuring that there were no air bubbles. Then 2 pieces of SSC soaked Whatman paper were added and flattened with a roller to ensure there are no air bubbles. The layers were covered with a large stack of hand-towel tissues and a weight was added. This was left to blot over-night. The following morning the membrane was retrieved from the blotting assembly and cross linked using the UV Stratalinker 2400 at ~254nm for 3 minutes, making sure that the side of the membrane

closest to the gel was face up. This was done to fix the protein-DNA complexes to the membrane.

2.4.8.6. LightShift Chemiluminescent EMSA Detection System

The LightShift Chemiluminescent EMSA Kit (Table 2.1) was used to detect protein-DNA complex binding. The manufacturers' instructions were optimised as follows:

The blocking buffer and wash buffer (4x) were incubated at 37°C until all the SDS in the solution was dissolved. The membrane was placed in a clean container and 30ml of blocking buffer was added. The container was rocked gently for 15 minutes. 75µl of stabilized Streptavidin-Horseradish Peroxidase was added to the buffer in the container and rocked for a further 15 minutes while 100ml of wash buffer dilution (1x) was prepared. The buffer/Streptavidin solution was decanted from the membrane. 20ml of wash buffer (1x) was added to the membrane and rocked gently for 5 minutes. This step was repeated 4 times in total. After decanting the last of the wash buffer the membrane was washed in 20ml of substrate equilibrium buffer for 5 minutes with gentle rocking. During this time 3ml of luminol/enhancer solution was combined with 3ml of stable peroxidase solution and decanted onto a clean glass plate. The membrane was then removed from the container and placed face down onto the solution on the glass plate, with no air bubbles, for 7 minutes. The membrane was briefly drip-dried and wrapped in Clingfilm, avoiding bubbles and wrinkles. It was then placed face up into a cassette for exposure to X-ray film.

2.4.8.7. EMSA Visualisation

To visualize the protein-DNA complexes the membrane was taken into a dark room and exposed to X-ray film for variable time points. The film was developed according to the manufacturers' instructions.

2.4.9. TaqMan Gene Expression

TaqMan Gene Expression was used to measure the amount of mRNA expression associated with a target sequence. To do this, complementary DNA (cDNA) of known genotype must be synthesised from RNA taken from lymphocytes.

2.4.9.1. RNA Extraction

10ml of whole blood was added to an equal volume of PBS (1x) (Table 2.1) and mixed by inversion. This blood/PBS mix was layered on top of 4ml Lymphocyte Separation solution (Table 2.1) and centrifuged at 2000rpm for 20 minutes. The resulting top layer

(plasma) was retained and added to 20ml PBS (1x) and centrifuged at 1000rpm for 5 minutes. Lymphocyte RNA were isolated using the RNeasy Mini Kit.

For the purposes of rs11573156 CC comparison with rs11573156 CG/GG samples, lymphocytes were isolated from 12 samples of 10ml whole blood from an in house European cohort known as the DRD cohort.

2.4.9.2. Complementary DNA (cDNA) Synthesis by Reverse Transcription

All incubations were carried out on the G Storm Thermocycler. For each reaction the following mix was prepared: 1µg RNA in a 10µl solution (made up with dH₂O), 2µl random primers (2 pMol) and 1µl dNTP (10mM). This mix was incubated at 65°C for 5 minutes and then the following was added: 4µl First Strand Buffer and 2µl DTT (0.1M). This was incubated at 42°C for 2 minutes and then 1µl Superscript III was added. After this addition the mix was incubated at 42°C for 40 minutes and then at 70°C for 15 minutes immediately after. Then 1µl ribonuclease (Rnase) H was added and the total mix was incubated at 37°C for 10 minutes. Samples were labelled and stored at -20°C.

2.4.9.3. TaqMan Gene Expression Assays (20x)

TaqMan Gene Expression Assays (20x) are target protein-coding transcripts. All assays used throughout this thesis were of human species and are detailed in Table 2.3. The assays used for the work throughout this thesis were Inventoried or Made-to-Order. These assays were pre-designed and tested. They were either mass manufactured and stored as stock or manufactured at the time of order.

TaqMan Endogenous Controls are a collection of predesigned assays for candidate control genes used in order to normalise for genotype specific differences in sample RNA. Endogenous controls were pre-run with all samples to determine the most appropriate control. The best endogenous controls show no difference in expression across all samples.

Each TaqMan Gene Expression Assay (20x) and Endogenous Control contains two unlabelled primers (18µM per primer) and one 6-FAM reporter dye labelled TaqMan MGB probe (5µM). In both instances probes have a reporter dye attached to the 5' end of the probe and a non-fluorescent quencher (NFQ) at the 3' end.

TaqMan Gene Expression Method Description and Conditions

ABI TaqMan Gene Expression reactions were carried out in a 384 well plate format. cDNA at 5ng/µl was transferred from a standardized array using the Biomek 2000 Laboratory Automation Workstation (Table 2.5) into a 384 well plate and centrifuged at

3000rpm for 1 minute. Each plate contained at least two no template control (NTC) wells, containing no cDNA, to control for contamination. A standard curve was produced for each assay using a serial dilution of cDNA at 0.25ng, 0.5ng, 1ng, 2ng and 5ng.

TaqMan gene expression reactions were carried out as per the TaqMan Gene Expression Mastermix manufacturers' instructions. Each plate was centrifuged at 3000rpm for 1 minute and PCR reactions were carried out using the 7900HT Fast Real-Time PCR system (Table 2.5). The standard PCR reaction conditions were as follows: Uracil-DNA Glycosylase (UDG) incubation at 50°C 2 for minutes then AmpliTaq Gold UP Enzyme Activation at 95°C for 10 minutes followed by 40 cycles of denaturing at 95°C for 15 seconds each and finally annealing/Extension at 60°C for 1 minute.

In a similar fashion to TaqMan Genotyping there are several phases that occurred during each PCR reaction for TaqMan Gene Expression (Figure 2.7). UDG was activated to prevent the re-amplification of carry over PCR-products. Each TaqMan MGB probe annealed to the specific complimentary sequence between the forward and reverse primer sites. At that point the probe was still intact and therefore the NFQ was still at close enough proximity to the reporter dye to quench its fluorescent signal. Next the AmpliTaq Gold DNA UP polymerase (a component of the TaqMan Gene Expression Master Mix) 5' to 3' activity along the DNA template between the forward and reverse bound primers specifically cleaved the probes that were hybridized to their target. The effect of cleavage was to separate the reporter dye from the NFQ, which caused an increase in fluorescence from the reporter dye and both probe fragments to become displaced from the target. Polymerization of the strand continued. The probes were blocked from the 3'end to prevent them from extending themselves during the PCR. During the PCR process the ROX passive reference acted as an internal reference for the normalization of the reporter dye signal, correcting for changes in concentration or volume.

The increase in fluorescent signal was only detected if the target sequence was complementary to the probe and amplified during PCR. This prevented non-specific amplification detection

(http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_039284.pdf).

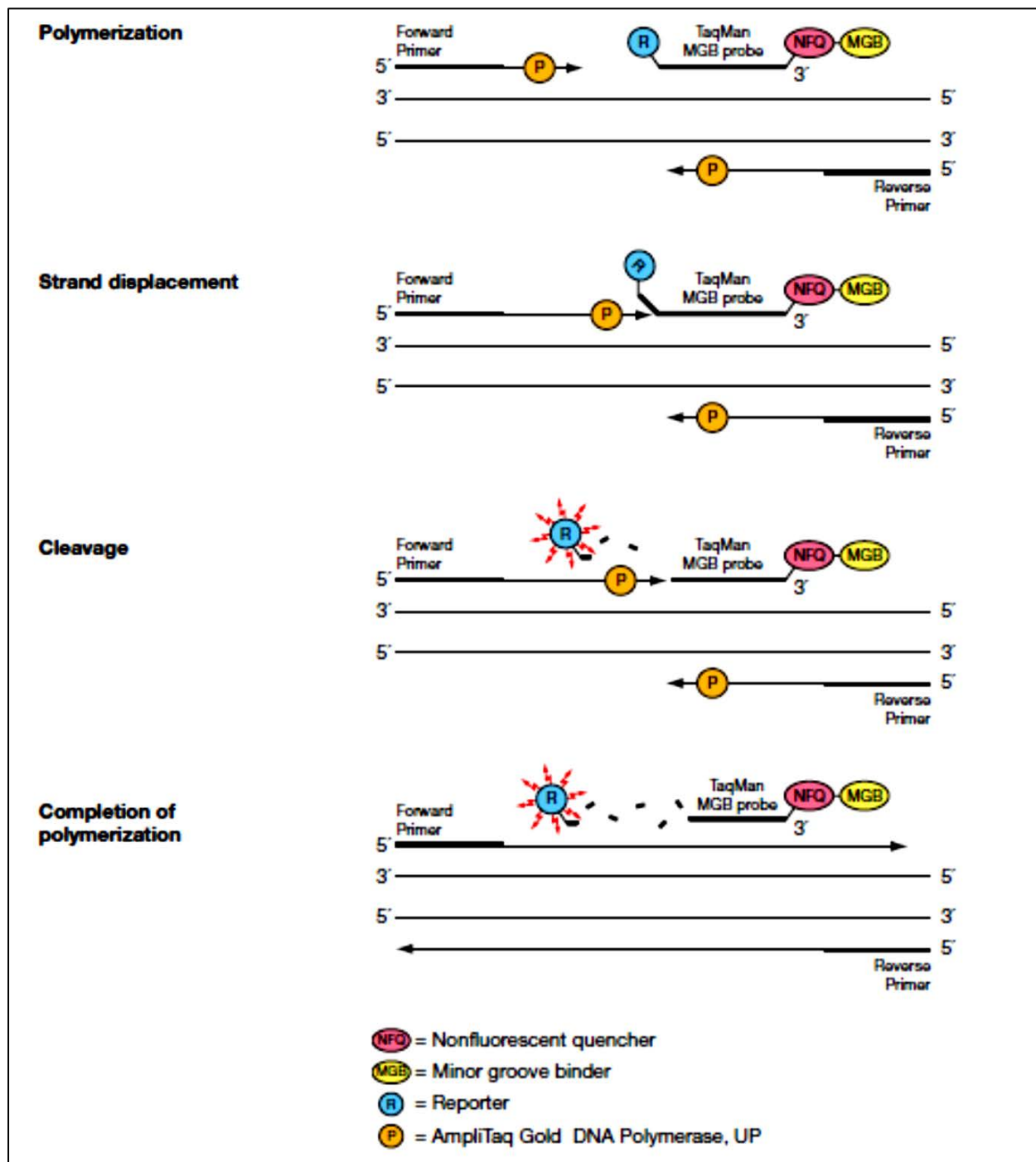


Figure 2.7. The 5' Nuclease Assay. Schematic to show the stages of a cycle of the TaqMan Gene Expression PCR . Reprinted from: http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_039284.pdf).

TaqMan Gene Expression Signal Detection

Analysis was carried out using the 7900HT Fast Real-Time PCR System (Table 2.5) with Sequence Detection System (SDS) v2.1 software (Table 2.4). Each well was manually labelled using the SDS software to measure the correct assay.

A relative quantitation analysis was used for all plates. This compares a target against an internal standard (ROX) using the comparative threshold cycle (Ct) method. Ct is the PCR amplification cycle number where the signal associated with a sample first shows a significant exponential growth due to increased fluorescent signal detection with amplification. The lower the Ct, the higher the expression. In addition all measurements were also normalized to endogenous controls, which accounted for the quality of the cDNA used and any concentration variability.

Normalization was automatically calculated using the SDS relative quantitation function. This was achieved by dividing the emission intensity of the reporter gene by the ROX passive reference at two time points. The first is at the baseline (the early cycles of PCR before significant exponential growth) (R_n^- value) and the second at 40 cycles - the PCR endpoint (R_n^+ value). The ΔR_n value is the difference between the R_n^+ and R_n^- values, which represents an accurate measurement of the magnitude of the signal generated by the PCR conditions (Figure 2.8).

The SDS software generated results as a set of amplification plots (Figure 2.9a) and a text file output. Plots can be manually viewed per sample or per assay (Figure 2.9b). This function allowed me to ensure that each assay had worked and to manually remove any anomalous results by discounting the relative well. Both baseline and threshold cycle values can be automatically calculated by the SDS software or set manually. For all assays carried out for rs11573156 I manually set the threshold to optimise the view of each set of results. The text output values were used to compare differences in expression by genotype. To achieve this, the text output values were sorted into genotype groups and input into the Rest-384 software system (Table 2.4). This calculated the average difference in ΔC_t values between genotypes in the target sequence and between genotypes in the endogenous controls, taking into account the standard curve values

(http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_039284.pdf).

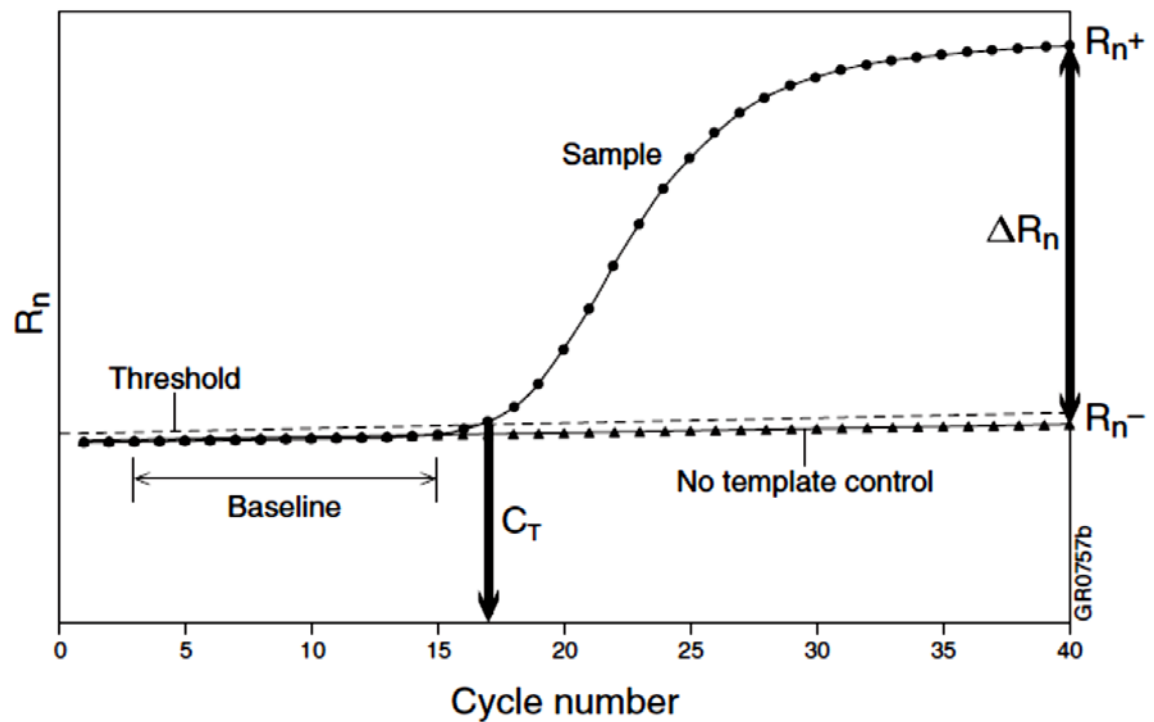


Figure 2.8. TaqMan Gene Expression Schematic to show how each set of results are interpreted. Normalization using ROX showing R_n as a function of PCR amplification cycle number. This graph represents how ΔR_n is calculated and also indicates the point of significant exponential growth that determines ΔC_T based on the increase in fluorescent signal detected with PCR amplification (http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_039284.pdf).

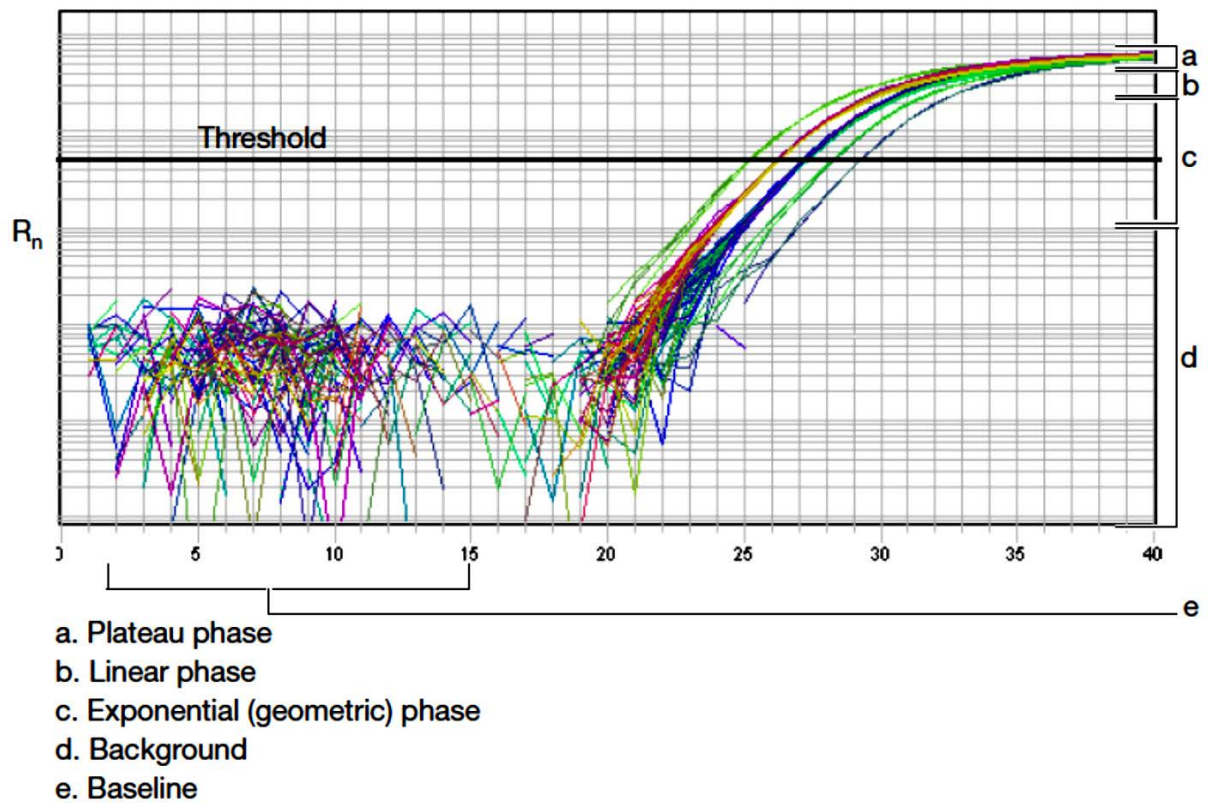


Figure 2.9a. A Typical Amplification Curve. This graph represents a typical amplification curve and identifies the main phases of the run. Each line represents an individual sample and each colour represents a particular target or endogenous control assay

(http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_039284.pdf).

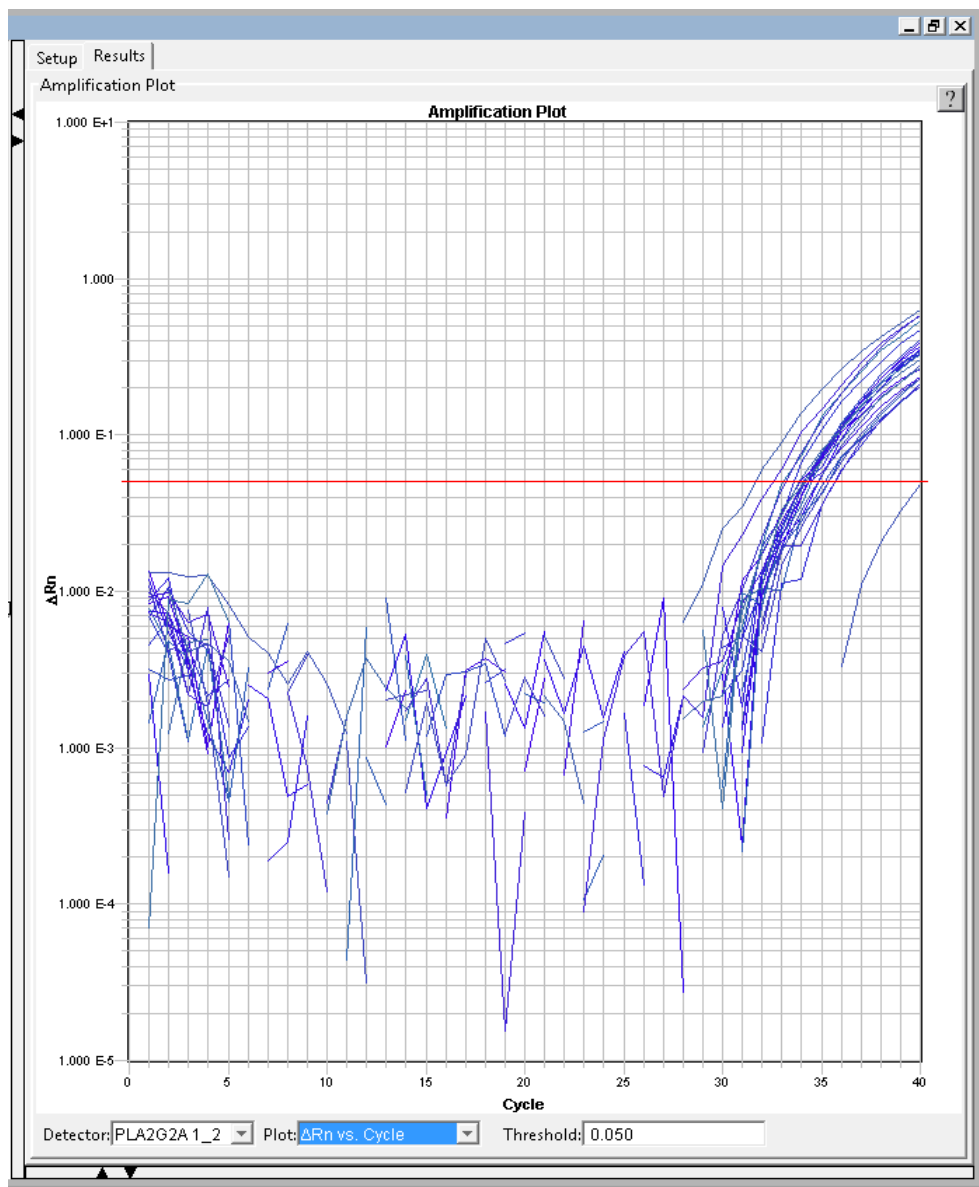


Figure 2.9b. This is an example of a typical results plot for *PLA2G2A* TaqMan gene expression using the SDS software. Unlike Figure 2.9a, which represents the results for each assay (each shown in a different colour) this figure has been set to show the results for the *PLA2G2A* exon 1-2 assay only. This allows one to interpret whether the assay has worked and to manually remove any anomalous results. For example the sample that does not reach the threshold before 40 cycles is anomalous and was removed from the analysis. The threshold was manually set in order to best visualise the amplification curve.

TaqMan Gene Expression Analysis for rs11573156 C>G

The effect of rs11573156 C>G on mRNA levels was assessed using a TaqMan gene expression assay in a set of 12 Caucasian samples from the DRD cohort: 5 samples were homozygous for the common allele (CC), there were 4 heterozygous samples (CG) and the frequency of homozygosity of the minor G allele (0.0625) meant it was only possible to

identify three GG homozygous samples from this volunteer cohort. The samples were taken from a total of 67 previously genotyped DRD study participants using the TaqMan Genotyping protocol (2.6.1) and analysed using the relative quantitation (using comparative Ct) program on the ABI Prism 7900HT TaqMan machine (Table 2.5). Only 12 samples could be utilised due to a lack of remaining blood stocks for many of the study participants.

RNA samples were isolated and reverse transcribed as described above for use with two TaqMan ABI inventoried gene expression assays (Table 2.3), Hs01044022_m1, spanning human *PLA2G2A* exon 1 and 2 and Hs00179898_m1, which spans exon 5 and 6. They were compared and normalised to the housekeeping genes Hs99999903_m1; ACTB (beta actin) and Hs99999905_m1; GAPDH. In all instances the housekeeping genes showed no difference in expression levels across all samples. Each sample was run in triplicate for each assay on the ABI Prism 7900HT Real Time PCR System. Results were analysed using the REST-384 Version 2 tool (Table 2.4).

2.4.10. The Advanced Study of Aortic Pathology (ASAP) Gene Expression Analysis

The Advanced Study of Aortic Pathology (ASAP) is a cohort founded by Prof. Per Eriksson and his team at the Karolinska University Hospital in Stockholm. All analysis relating to this study was kindly carried out on my behalf by a member of the Karolinska team, Dr. Lasse Folkersen.

The Advanced Study of Aortic Pathology (ASAP) recruited 221 patients undergoing aortic valve surgery at the Karolinska University Hospital, Stockholm Sweden (Sitia, Tomasoni et al. 2010). Tissue biopsies were taken from liver, mammary arteries and dilated and non-dilated ascending aorta and heart during surgery. The medial and adventitial layers of the vascular specimen were isolated by adventicectomy, and incubated with RNAlater (Ambion, Austin, Texas, USA) and homogenised for mRNA extraction (Sitia, Tomasoni et al. 2010). Affymetrix Gene Chip Human Exon 1.0 ST expression arrays were used. Messenger-RNA expression was evaluated using RNA pre-processing (Luscher and Barton 1997). Participants were genotyped using the Illumina Human 610W-Quad Bead array (Ross 1999), which included 101 SNPs in the region 200kb up and downstream from the *PLA2G2A* locus. SNPs not covered by the Illumina Human 610W-Quad Bead array were imputed for the 1000 genome project. Classification by mach1 suggests >0.3 as the lowest threshold for R^2 . In our set this corresponded to ~10% mistakes. Per allele mRNA expression levels were calculated on an individual SNP basis for SNPs of interest that were available on the array.

2.4.11. Statistics

Specific statistical analysis is indicated as appropriate, along with clear acknowledgements of who has carried out each analysis. In general the programs used throughout the project included;

Stata Data Analysis and Statistical Software, which was mainly used for comparing genotype data to assess the effect of different genotypes on sPLA2 levels and activity, cardiac events and cardiac related disorders.

IBM SPSS Statistics, which was mainly used to compare genotype effects on luciferase activity using results from the luciferase assays. More specifically, statistical analysis of luciferase assays was performed using SPSS version 12.0 and an independent sample T-test was run to determine the P-value, with a threshold of $P < 0.05$.

For the rs11573156 Mendelian randomisation a pre-specified Stata script was developed by Jackie Cooper and Michael Holmes to standardise the genetic analysis in the collaborating studies. The genotyping was arranged so that the genotype coding was directionally consistent and reduced sPLA2 with the addition of the C allele. Traits considered not to be normally distributed included sPLA2-IIA mass, sPLA2 activity, carotid intima medial thickness (C-IMT), C-reactive protein (CRP), triglyceride and interleukin-6 (IL-6) concentrations. These traits were log(n) transformed and differences between genotype groups were reported as a relative or percentage difference. In order to combat the potential for established vascular disease altering the magnitude of genotype-trait association, there was limited analysis of associations between genotype and continuous traits to controls in retrospective case-control studies. Regarding cohort studies and nested case-control studies, all participants were included at recruitment, as variable measurements were made prior to clinical events.

For the *PLA2G2A* SNP rs3767221 and the *PLA2G5* and *PLA2G10* meta-analyses statistical analysis was carried out by Jackie Cooper and Michael Holmes. A pre-specified Stata script was developed to standardise the genetic analysis in the collaborating studies. A trait considered not to be normally distributed was sPLA2 activity. This trait was log(n) transformed and differences between genotype groups were reported as a relative or percentage difference. In order to combat the potential for established vascular disease altering the magnitude of genotype-trait association, there was limited analysis of associations between genotype and continuous traits to controls in the retrospective case-control study. Regarding cohort studies and nested case-control studies, all participants were included at recruitment as variable measurements were made prior to clinical events.

Regarding the genotyping results for *PLA2G10* SNPs of interest, statistical analysis of RFLP and TaqMan genotyping data was analysed by Jackie Cooper using Stata.

For the ASAP study statistical analysis of microarray expression data was performed using R 2.13.0 (Folkersen, Diez et al. 2009). P-values are calculated using a linear additive model encoding genotypes numerically. Plotting of location dependent associations was performed using the Gene Region Scanversion 1.8.0 (Libby 1995).

2.5. Meta Analyses Study Populations and Specifications

All of the studies populations analysed throughout this thesis correspond to samples from European populations. Informed consent was obtained from all participants and permissions were granted by all relevant governing bodies where necessary. All additional population information is shown in Tables 2.6-2.7.

2.5.1. Mendelian Randomisation Using *PLA2G2A* rs11573156

To carry out the Mendelian randomisation detailed in **Chapter 4**, we first had to consider how many individuals to include in the meta-analysis to give sufficient power. A power calculation was carried out to determine that the numbers involved in the Mendelian randomisation were considered great enough to determine a conclusion with high accuracy and reliability: The power calculation is described as follows:

To estimate the power, we used the gene effect on sPLA2 (the per-allele effect was to reduce sPLA2-IIA by approximately 0.5 SD) and the observational association (a 1SD reduction in sPLA2-IIA mass associated with an OR of CHD of 0.84), we estimated the expected effect assuming the observational association was causal - i.e. the per-allele effect of sPLA2-IIA should associate with an OR of $0.84^{0.5}=0.92$. Assuming a 10% risk of CVD in non-carriers of the variant, the risk would be 9.2% in individuals carrying a variant (assuming a causal relationship). We estimated that we would need 15,705 cases in 157,050 to have 90% power at $\alpha=0.05$.

Using the values that we have: 19,716 cases (combined total of incident, prevalent and recurrent) and 96,523 controls we have a beta value = 0.07 and the power = $1 - \beta = 0.93$ (93%), where beta value is the likelihood that an experiment will yield a Type II error.

Our collaboration incorporated data from 32 studies of 96,523 individuals of European descent. Of these studies, 22 were considered to be in general populations, including 15 cohorts, 3 nested case-control and 4 case-control studies. There were also 8 studies carried out in patients with acute coronary syndrome (ACS); 7 cohorts and 1 nested

case-control study. A cohort of patients with established vascular disease or risk factors for CVD (SMART) and a case-control study of ACS patients undergoing percutaneous coronary intervention (PCI) in which the outcome was coronary artery restenosis (GENDER). One additional study to these 32, the ASAP study, was not included in the meta-analysis, but was used to analyse the *PLA2G2A* mRNA expression associated with rs11573156. In all instances approval from relevant ethical committees was obtained and all analyses unless otherwise stated was performed using Stata and led by Dr Michael Holmes. Twenty-eight of these studies were used in the final Mendelian randomisation analysis. Table 2.6 gives an overview of the studies included and Table 2.7 indicates the relevant CHD and sPLA2 associated measurements that could be obtained from each study.

Study	Study design	Geographical location	Sampling frame	Participants included	Baseline year(s)	% Female	Age, mean (SD)
STUDIES IN GENERAL POPULATION							
AMC-PAS (Langer and Gawaz 2008)	C-C	Netherlands	Hospital clinic	740	1990-2000	19.00	44.05(3.91)
BHF-FHS (Lusis 2000)	C-C	UK	Direct media campaign	4521	1998-2003	36.67	50.53(13.90)
BRHS (Zhang 2008)	Cohort	UK	General practice	3835	1998-2000	0.00	68.74(5.49)
BWHHS (Zhang 2008)	Cohort	UK	General practice	3405	1999-2001	100.00	68.91(5.49)
CCHS (Steinberg 1987)	Cohort	Denmark	General population	10375	1991-1994	55.68	56.57(16.26)
CYPRUS (Lusis, Mar et al. 2004)	Cohort	Cyprus	Mayor's list	734	2003-2008	52.97	61.27(10.26)
EAS (Brown and Goldstein 1986)	Cohort	UK	General practices	857	1987	50.64	64.38(5.82)
EPIC-Netherlands (Brown and Goldstein 1997)	Nested C-C	Netherlands	Existing cohorts	5194	1993-1997	78.11	54.03(10.22)
EPIC-Norfolk (Lusis 2000)	Nested C-C	UK	General practices	3039	1993-1997	34.65	64.68(7.88)
GRAPHIC (Berliner and Heinecke 1996)	Cohort	UK	Nuclear families	2024	2003-2005	49.56	39.19(14.48)
IMPROVE (Badimon, Storey et al. 2011)	Cohort	Europe	Clinic	3236	2004-2005	52.83	64.35(5.17)
LIFE Heart (Pentikainen, Oorni et al. 2000)	C-C	Germany	Hospital clinic	3128	2006-present	32.83	62.86 (11.52)
MedStar (Orso, Grandl et al. 2011)	Cohort	USA	Single-centre hospital	1322	2004-2007	54.59	59.76(8.86)
NPHS-II (Orso, Grandl et al. 2011)	Cohort	UK	General practices	2693	1989-1994	0.00	56.11(3.33)
PennCath (Libby 1995)	C-C	USA	Single-centre hospital	1516	1998-2003	51.92	59.29(9.69)
PREVEND (Borissoff, Spronk et al. 2011)	Cohort	Netherlands	Community	8114	1997	50.86	49.05(12.76)
PROSPER (Ross 1999)	RCT	UK, Ireland, Netherlands	General practices	3991	1997-1999	51.79	75.28 (3.35)
Rotterdam (Prentice, Paczesny et al. 2010)	Cohort	Netherlands	District participants	5974	1989-1993	59.37	69.38(8.91)
TPT (Graham 2006)	RCT	UK	General practice	4014	1984-1989	0.00	56.06(6.74)
UCP (Duscha, Slentz et al. 2005)	Nested C-C	Netherlands	National drug registry	1632	1985-2005	25.49	62.76 (9.64)
UDACS (Bhupathiraju and Tucker 2011)	Cohort	UK	Clinic	564	2001-2002	41.13	66.73(11.09)
Whitehall II (Ockene and Miller 1997)	Cohort	UK	Workplace	5018	1985-1988	26.44	43.90(5.93)
STUDIES IN ACUTE CORONARY SYNDROME PATIENTS							
FAST-MI (Doll and Hill 1966)	Cohort	France	Nationwide ACS registry	973	2005-ongoing	29.19	66.09(13.65)
GENDEMIP (Gordon, Kannel et al. 1974)	Cohort	Czech Republic	5 Coronary Units	1432	2006-2009	25.98	57.08(8.62)
GRACE-France (2004)	Cohort	France	Hospitals	274	2000-2002	20.44	60.13(12.67)
GRACE-Scotland (Ambrose and Barua 2004)	Cohort	Scotland	Hospital	1488	1999-2009	30.04	64.84 (12.04)
IHCS (Hawe, Talmud et al. 2003)	Nested C-C	USA	Single hospital	2382	1994-ongoing	29.14	62.99 (12.29)
KAROLA (Boer, Feskens et al. 1999)	Cohort	Germany	Rehabilitation clinics	1019	1999-2000	15.00	58.93 (7.96)
MERLIN-TIMI 36 (Pohjola-Sintonen, Rissanen et al. 1998)	RCT	17 countries	Hospitals	1606	2004-2007	34.10	63.43 (10.81)
PROVE-IT TIMI 22 (Willer, Sanna et al. 2008)	RCT	8 countries	Hospitals	2260	2000-2003	22.51	57.48 (11.06)
OTHER STUDIES							
GENDER (Boer, Feskens et al. 1999)	C-C	Netherlands	Hospital clinic	866	1999-2001	26.00	62.01 (11.09)
SMART (Hawe, Talmud et al. 2003)	Cohort	Netherlands	Hospital clinic	8297	1996-1998	32.24	56.51(12.42)
ASAP (Pohjola-Sintonen, Rissanen et al. 1998)	Cohort	Sweden	Hospital clinic	272	2006-2010	28.68	63.54 (12.39)

Footnotes: C-C: case control; ICU: intensive care unit; RCT: randomized clinical trial; X-S: cross sectional. All cohorts were prospective in design.

Table 2.6. The studies included in the MR analysis of rs11573156 with sPLA2-IIA mass/sPLA2 activity and the risk of CHD. The table includes details of the study backgrounds.

Study	sPLA ₂ assays		Circulating biomarkers/anthropometric traits								
	IIA mass	Activity	SBP	LDL-C	HDL-C	TG	IL6	CRP	Glucose	BMI	C-IMT
STUDIES IN GENERAL POPULATIONS											
AMC-PAS			•	•	•	•		•	•	•	
BFH-FHS											
BRHS			•	•	•	•	•	•	•	•	
BWHHS			•	•	•	•	•	•	•	•	
CCHS			•	•	•	•		•	•	•	
CYPRUS			•	•	•	•		•		•	•
EAS			•	•	•	•	•	•	•	•	•
EPIC-Netherlands			•	•	•	•		•	•	•	
EPIC-Norfolk	•	•	•	•	•	•		•		•	
GRAPHIC			•	•	•	•		•	•	•	
IMPROVE			•	•	•	•		•	•	•	•
LIFE Heart				•	•	•		•	•	•	•
MedStar			•	•	•					•	
NPHS-II			•	•	•	•		•		•	
PennCath			•	•	•					•	
PREVEND			•	•	•	•		•	•	•	•
PROSPER			•	•	•	•	•			•	
Rotterdam			•		•		•	•		•	•
TPT			•							•	
UCP										•	
UDACS	•		•	•	•	•	•	•	•	•	
Whitehall II			•	•	•	•	•	•	•	•	
STUDIES IN ACUTE CORONARY SYNDROME POPULATIONS											
GENDEMIP			•	•	•	•		•	•	•	
FAST-MI	•	•	•				•	•	•	•	
GRACE-France	•	•	•	•				•		•	
GRACE-Scotland	•	•	•	•	•	•		•	•	•	
IHCS			•	•	•	•		•	•	•	
KAROLA	• †	•	•	•	•	•	•	•	•	•	
MERLIN-TIMI 36											
PROVE-IT TIMI 22											
OTHER STUDIES											
GENDER											
SMART											•

Footnotes: BMI: body mass index; C-IMT: carotid intima medial thickness; CRP: C-reactive protein; HDL-C: high-density lipoprotein cholesterol; IL6: interleukin 6; LDL-C: low-density lipoprotein cholesterol; SBP: systolic blood pressure; sPLA₂: secretory phospholipase A2; TG: triglycerides; † not included in the *PLA2G2A*-sPLA2 analysis as blood samples were not taken at time of ACS. All measurements undertaken by individual study groups and have been previously reported in published data.

Table 2.7. This table represents the available CHD/sPLA2 relevant biomarker measurements for each included study.

Mendelian Randomisation Method

Data was taken from 28 of 32 studies for the analysis of the association between rs11573156 and putative and established cardiovascular risk factors; 21 general population studies, 6 ACS cohorts and 1 established vascular disease cohort. This data was used to test the associations of *PLA2G2A* with cardiovascular risk and other co-variables. For these continuous traits, per C allele associations of rs11573156 with variables were estimated within each study by linear regression analysis. The studies not included here were BFH-FHS, MERLIN, PROVE-IT, GENDAR and SMART.

When analysing the associations between rs11573156 and cardiovascular events, data was taken from 29 studies including 21 general population studies and 8 ACS patient studies. Two additional studies, SMART and GENDER, were also analysed, but could not be included due to their patient characteristics or outcome definitions differing from the other studies.

The cardiovascular outcomes were defined as follows: (i) Major Vascular Event (MVE) (comprising fatal/non-fatal myocardial infarction or stroke); (ii) non-fatal myocardial infarction (ST-segment elevation myocardial infarction or non-ST-segment elevation myocardial infarction); (iii) non-fatal stroke (including ischemic or haemorrhagic subtypes), and; (iv) fatal myocardial infarction or stroke.

Studies in general populations were divided into prevalent (events occurring pre-recruitment) or incident (events occurring during follow-up in individuals previously free from MVE). For all ACS studies, all events after recruitment were included and labelled as recurrent events. The within-study per C allele association of rs11573156 with MVE was estimated using logistic regression.

For all continuous traits, the results were pooled using a fixed effect (inverse variance) meta-analysis and for all clinical traits estimates were pooled using fixed-and random-effects meta-analysis. All meta-analysis was stratified by clinical setting and used I^2 to measure heterogeneity (Marenberg, Risch et al. 1994). Assuming the association between sPLA2-IIA mass and fatal/non-fatal MI followed a log-linear relationship and was free from bias, we estimated the expected effect of the genetic variant on risk of MVE, and compared this to that observed outcome.

PLA2G5 Study Populations Used to Analyse SNPs of Interest

Based on the initial results from UDACS, I and others in our lab genotyped the selected SNPs in up to 11 additional studies, which are listed and detailed in Table 2.8.

There was also limited SNP genotyping data available from MetaboChip genotyping. The MetaboChip is a custom Illumina iSelect genotyping array designed to test ~200,000 SNPs of interest for disease phenotypes such as; metabolic disease and atherosclerotic/cardiovascular disease. The content of the chip was chosen based on large scale meta-analysis of relevant traits (including up to 100,000 individuals) and of HapMap and 1000 Genomes Project SNP content (<http://www.sph.umich.edu/csg/kang/MetaboChip/>). Several MetaboChip studies agreed to contribute genotype data.

Using the genotyping data available, associations for the identified SNPs with CHD risk and known CHD risk factors was studied and additionally the associations between these SNPs and sPLA2 activity were studied. We also have access to data from the GWAS study; Wellcome Trust Case Control Consortium (WTCCC), thanks to Dr. Reecha Sofat who is a registered user for WTCCC data. WTCCC is a UK based GWAS, established in 2005 (Lewington, Clarke et al. 2002).

Study:	SNPs Genotyped in Each Study:				Genotyped By:
	rs11573185	rs11573248	rs11573203	rs11573191	
UDACS (Mason 2011)	•	•	•	•	P. Wootton
EPIC-Norfolk (Jones, Dhamrait et al. 2003)	•	•		•	M. Guardiola
GRACE-FRANCE (Kuller and Meilahn 1996)	•		•	•	M. Guardiola
Edinburgh Artery Study (Willerson and Ridker 2004)	•	•			H.Exeter
NPHSII (Corson 2009)	•	•	•	•	K.Li
TPT (Andreotti, Porto et al. 2002)	•	•			H.Exeter
BWHHS (Jain and Ridker 2005)	•				MetaboChip
Whitehall II (Ray, Stein et al. 2002)	•				MetaboChip
Women's Health Initiative (Kimmel, Berlin et al. 2004)	•				MetaboChip
EPIC-Netherlands (Ali, Narayan et al. 2010)	•				MetaboChip
IMPROVE (Kannel and McGee 1979)		•			H.Exeter
CYPRUS (Moss, Klein et al. 1991)		•			H.Exeter

Footnotes: BWHHS: British Women's Health and Heat Study; Cyprus: Cyprus Study; EPIC: European Prospective Investigation into Cancer and Nutrition; GRACE: Global Registry of Acute Coronary Events; IMPROVE: IMPROVE Study; NPHS-II: Northwick park Heart Study; TPT: Thrombosis Prevention Trial; UDACS: University College London Diabetes And Cardiovascular Disease Study; Whitehall II: Whitehall II Study.

Table 2.8. A list of the *PLA2G5* tSNPs identified by univariate analysis in UDACS as having potentially significant associations to either total cholesterol or LDL levels (CHD risk factors) or sPLA2-IIA mass and the studies that each SNP has been genotyped in. The cohort details for these studies have been detailed previously in Table 2.6 and the biomarkers measured in each study have been previously described in Table 2.7.

PLA2G10 Study Populations Used to Analyse SNPs of Interest

The studies included in the genotyping analysis are listed in Table 2.9, which also gives an overview of the traits measured in each study. Initially all four SNPs were genotyped in EPIC-Norfolk and GRACE and, based on these results (6.3.3), rs4003228 was additionally genotyped in a further two studies.

SNPs Genotyped				Genotyped By:
rs4003232	rs4003228	rs72546339	rs72546340	
STUDIES IN GENERAL POPULATIONS				
EPIC-Norfolk (n=3039)	●	●	●	H.Exeter/M.Gardiola
NPHS-II (n=2693)	●			H.Exeter/M.Gardiola
UDACS (n=564)	●			H.Exeter/M.Gardiola
STUDIES IN ACUTE CORONARY SYNDROME POPULATIONS				
GRACE-France (n=274)	●	●	●	H.Exeter/ M.Gardiola

Table 2.9. The baseline characteristics for EPIC-Norfolk, GRACE-France, UDACS and NPHS-II. The cohort details for these studies have been detailed previously in Table 2.6 and the biomarkers measured in each study have been previously described in Table 2.7. This table shows which of the previously identified *PLA2G10* tagging SNPs were genotyped in which of the included studies and who was responsible for the genotyping.

The DRD Cohort

The DRD cohort is an in-house volunteer cohort. A small subset of volunteers who work at The Rayne Building or live in the surrounding area of London in the UK were invited to donate a blood sample for genetic analysis. In all cases baseline data of age at donation and gender were taken. Gender was approximately 50% male and 50% female and ages ranged between 19 and 65 years old.

CHAPTER 3

Identification of Functional Variants of *PLA2G2A*

3. Background

Secretory PLA2-IIA has been identified as a novel biomarker of CHD in previous observational studies (Boekholdt, Keller et al. 2005; Mallat, Steg et al. 2005; Mallat, Benessiano et al. 2007; Dutour, Achard et al. 2010) and it has been suggested that this is potentially due to its known inflammatory effects playing a role in atherosclerosis and inflammation related disease (**1.5.7**). In order to verify the role of sPLA2-IIA as a causal biomarker for atherosclerosis, genetic analysis can be employed. Identification of genetic variants that are significantly associated to sPLA2-IIA could be used as genetic instruments to test the causality of sPLA2-IIA for atherosclerosis using Mendelian randomisation. This approach can be used as a proxy for a randomised control trial (RCT); negating the need to find a novel drug against the biomarker and a cohort of patients to test it on before causality of the biomarker is established, saving both time and resources.

3.1. Introduction

Chapter **1.7.2** describes the identification of six tSNPs across the *PLA2G2A* gene. Of these six SNPs two are shown to have a significant effect on sPLA2-IIA levels; rs11573156 C>G and rs3767221 A>C. The aim of this chapter was to determine if these two SNPs are functional and contributing directly to the regulation of levels of sPLA2-IIA, or whether they are merely marking a functional variant elsewhere in the gene. Identification of functional variants would determine viable genetic instruments for Mendelian randomisation and give a greater depth of understanding to the mechanism of sPLA2-IIA.

For both rs11573156 and rs3767221, research into the characteristics and possible effects of these SNPs was carried out using several freely available bioinformatics sources (**2.2**). In order to confirm whether a SNP is likely to be functional it is beneficial to use these bioinformatics tools to obtain known and predicted data about a SNP of interest. This data can relate to a huge number of things including; SNP type (missense, synonymous or nonsense), SNP position, known and potential transcription factor binding sites, open chromatin regions, known and predicted alternate splicing sites, other SNPs in LD, SNP conservation, allele frequency and potential microRNA (miRNA) binding sites. Prediction algorithms such as MatInspector, which predicts transcription factor binding sites and MiRBase, PITA, MiRTar and RegRNA, which predict miRNA binding sites were also used.

Transcription factor binding usually occurs in the promoter region and distal parts of the 5'UTR (Chatterjee and Pal 2009). Transcription factor proteins binding to a promoter can behave as enhancers or repressors to regulate the transcription of a gene. Since the promoter region of *PLA2G2A* has not been conclusively defined, it is possible that rs11573156 may affect a transcription factor binding site and it was therefore prudent to use the MatInspector tool to investigate this. Micro-RNA binding usually occurs in the 3'UTR and regulates the sequence post-transcriptionally, potentially enhancing or repressing levels of protein expression. However, there are instances where miRNAs have been shown to bind in the 5'UTR and contribute to post-transcriptional regulation (Tsai, Lin et al. 2009). Prediction algorithms for miRNA binding were therefore used to examine the regions surrounding both rs11573156 (5'UTR) and rs3767221 (3'UTR).

Collating bioinformatics data, both known and predicted, about a SNP can allow us to predict its potential as a functional SNP and may suggest mechanisms through which the SNP could act. This information may help to direct how the investigation of a SNP will progress and which techniques will be the most practical for analysing if and how a SNP functions to contribute to the actions of the gene and ultimately the protein produced.

In order to confirm the potential functionality of rs11573156 and rs3767221 for changes in sPLA2-IIA levels and/or sPLA2 activity I employed several techniques. The Dual-Reporter Luciferase assay (Table 2.1 and **2.4.7**) was used in order to measure the change in luciferase activity between the common and rare variants of the SNPs of interest. The change in luciferase activity is taken as an *in vitro* replication of the ability of each SNP to affect the levels of transcription of the gene. A significant difference in luciferase activity may indicate that a SNP is important in the regulation of gene transcription and therefore the levels of biomarker produced by the gene. Therefore, genotyping an individual for this SNP would suggest whether a person was likely to have increased or decreased sPLA2-levels.

Often changes in transcription levels associated with a SNP are due to the change in allele altering a transcription factor binding site. Transcription factor proteins bind to DNA at specific, recognisable binding sequences and help to regulate transcription. If one variant of a SNP is part of a transcription factor binding site, then the other variant of the SNP will cause the transcription factor binding site to become unrecognisable and the transcription factor will not bind there. EMSA (Table 2.1 and **2.4.8**) analysis can be used to detect protein binding to a sequence of interest. By isolating a sequence fragment for both the wild type and rare alleles of a SNP of interest, I was able to use EMSA analysis to compare protein binding between the two alleles. This can be used to support a null difference in luciferase assay results between two alleles, by showing no differential protein binding between the

same two alleles. However, should the sequence containing one allele have a protein binding site not present within the sequence containing the second allele it is likely that there is differential transcription factor binding between these variants. Using the bioinformatics results for a SNP that shows differential protein binding and taking into account the direction of change in luciferase assays between each allele (whether allele with the additional protein binding site showed increased or decreased luciferase activity) we can predict potential transcription factors that might be differentially binding to the SNP. A supershift EMSA uses a known antibody to the suspected binding protein to establish the identity of the differentially binding regulatory protein (Tsai, Smider et al. 2012).

If differential luciferase activity cannot be associated with a SNP and there is no visible difference in protein binding between alleles, it may be that the SNP is acting to regulate a gene post-transcriptionally. Post-transcriptional modifications may include miRNA binding, which can be predicted by bioinformatics algorithms (Table 2.4 and **2.4.1**). It may also include alternate splicing of a sequence including a SNP of interest, which would not affect transcription levels, but may affect levels of mRNA. In order to determine whether this is true for rs11573156 we carried out TaqMan gene expression assays (Table 2.1 and **2.4.9**) on samples of known genotype to determine any differences in expression associated with the SNP. This was compared to TaqMan assay results of the same genotyped samples in an area of the gene not considered to be differentially spliced. If a significant difference in expression was associated with the SNP of interest in the suspected splice region but no difference was seen between the same groups in a different part of the gene, it is likely that there is differential splicing in the region of the SNP. To support this data, further mRNA expression data was obtained from expression quantitative trait loci (eQTL) studies. These genomic loci regulate expression levels of mRNA or proteins. They are mapped using standard QTL mapping techniques that test the linkage between variation in expression and a genetic polymorphism. In this chapter eQTL data from the ASAP study (described in detail in **2.4.10**) was used to investigate the potential differential mRNA levels associated with our SNP of interest, rs11573156 in several different tissues.

3.2. Results

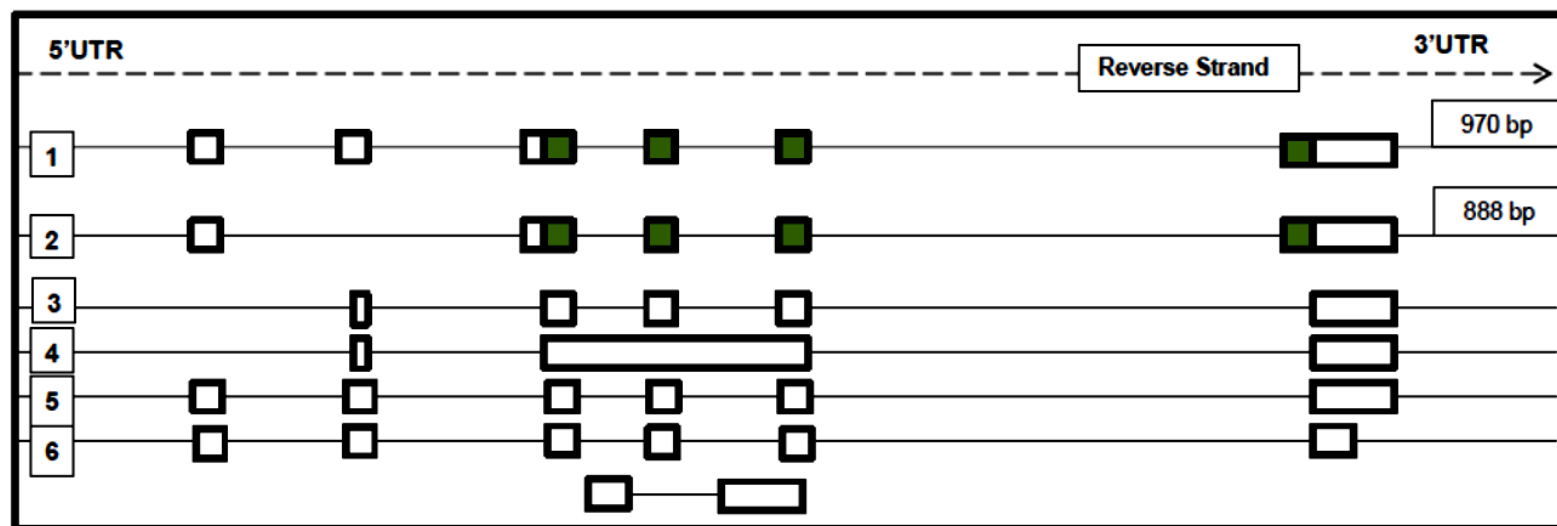
3.2.1. Bioinformatics Study of *PLA2G2A*

PLA2G2A exists in a cluster of *PLA2* genes on chromosome 1 (Figure 1.11). The gene is usually reverse transcribed and produces the enzyme sPLA2-IIA. The Ensembl Genome Browser (release 69) states that *PLA2G2A* is a known protein coding gene since it has been shown to be identical to known cDNAs or proteins from the same species. According to the Ensembl Genome Browser and NCBI databases there are 7 known

transcripts of the gene and two of these are described as known protein coding (Figure 3.1). These transcripts have been confirmed as protein coding since they have been shown to be 100% identical to an entry in either the Reference Sequence (RefSeq NP) database or the Swiss-Port database. The other 5 are described as 'processed transcripts,' which lack open reading frames and have no protein product (Figure 3.1). This data was collated by manual annotation carried out by the Havana group at the Wellcome Trust (http://vega.sanger.ac.uk/info/about/man_annotation.html). One of the known protein coding transcripts is made up of 5 exons (spliced) and the other of 6 exons (full length). However, both transcripts code for a 144 amino acid protein because in both transcripts only 4 exons are protein coding, the same 4 exons in both instances. The NCBI database includes the RefSeq collection, which is the result of data extraction from International Nucleotide Sequence Database Collection (INSDC) submissions, curation, and computation, as well as extensive collaboration with authoritative groups. RefSeq quotes the predominant version as NCBI ID: NM_000300.3, the longer (6 exon) transcript and NM_001161728.1 as missing one of the non-coding 5'UTR exon compared to the dominant transcript, but coding for the same protein. This is based on protein sequences derived from translation of the coding sequences submitted to public nucleic acid databases. A critical review of all data and peer reviewed literature is then carried out in order to annotate these transcripts. In Figure 3.1 the coding exons are shown as filled blocks and the non-coding exons are shown as empty boxes.

The existence of two protein coding transcripts coding for the same 144 amino acid protein (sPLA2-IIA), but different mRNAs, suggests alternate splicing of a non-coding exon. This may be relevant when investigating the impact of genetic factors on post-transcriptional regulation of the levels of sPLA2-IIA produced, depending on whether the transcripts are shown to be associated to a genetic polymorphism.

PLA2G2A Transcripts



5'UTR (Including promoter): Chr1:20307204

3'UTR : Chr1: 20301661

Green Box : Translated exon

White Box : Untranslated exon

Protein Coded By 1 and 2 : sPLA2-IIA;

144 amino acids long

1. Known protein coding transcript – shown by cDNA cloning and thought to be the pre-dominant form

2. Known protein coding transcript – shown by cDNA cloning. Missing non-coding exon 2 (82bp) compared to 1

3-5. Known processed transcript- predicted by manual annotation on case-by-case analysis (Havana Group). These transcripts are non-protein coding and do not contain an open reading frame.

Figure 3.1. *PLA2G2A* Transcripts. This schematic is based on data from the Ensembl Genome Browser (Release 69) and shows the known transcripts for the *PLA2G2A* gene. Of these, only two transcripts are protein coding. Five have been shown to be non-coding transcripts and classified as not containing an open reading frame (Though Ensembl does not specify the length of amino acids that 'no open reading frame' constitutes).

3.2.2. Bioinformatics Study of rs11573156

Data from NCBI and Ensembl confirms that, for the full length *PLA2G2A* transcript, rs11573156 is positioned at chromosome 1:20,306,146, which is in the untranslated exon 2, part of the 5'UTR and suspected promoter region of *PLA2G2A*. The SNP is located 38bp from the intron 1-2/exon 2 boundary in the full length transcript. In the alternately spliced protein coding transcript this SNP is spliced out. The European genotyping data available from HapMap and the 1000 Genomes Project showed that, on the reverse strand of the gene, the common allele, C occurs at a frequency of 0.75 and the rare G allele occurs at a frequency of 0.25. This SNP is a tagging-SNP in strong linkage disequilibrium (LD) with several other *PLA2G2A* SNPs; With an R^2 above 0.7 these SNPs are; rs4744 (synonymous coding SNP located at 1:20,304,926), rs2307246 (1:20,304,857, intron 1-2) and rs10732279 (1:20,305,065, intron 2-3). These LD values were confirmed in both 1000 Genomes Phase 1 CEU population (Utah residents with Northern and Western European ancestry) and the HapMap Project Phase 3 CEU population (again, Utah residents with Northern and Western European ancestry), a total population of ~1,500 individuals. Figure 3.2 is an LD heatmap using data from the HapMap Release #28 CEU population for *PLA2G2A* chr:20,301,940–20,306,426.



Figure 3.2. An LD heatmap showing the D' linkage disequilibrium between several SNPs of *PLA2G2A* (Chr1: 20,301,940–20,306,426) including rs11573156, rs4744, rs10732279 and rs2307246 (boxed in green). Data was taken from the HapMap Project Release #28 and the heatmap was created using Haploview version 4.0.

According to the MatInspector transcription factor binding prediction tool, there were five potential differential transcription factor (TF) binding sites at rs11573156 C>G. This was based on a 240bp FASTA sequence taken from Ensembl Genome Browser and representing the reverse strand of *PLA2G2A*. This sequence included rs11573156 at position 124. However, the prediction algorithm generated no potential differential TF binding site with an anchor site within 5 bases (position 119-129) of the SNP of interest at a 0.95 or above matrix similarity (the calculation of similarity between the sequence and the known TF binding site sequence). It is therefore unlikely that there is a differential transcription factor binding in this region. The only potential differential TF site (matrix similarity = 0.899) was a GC-Box factor Sp1/GC site that was present with the rare G allele but not with the common C allele and EMSA experiments were carried out to rule this out as a potential differentially binding protein (**3.3.5.1**).

MatInspector data comparing rs11573156 C>G is shown below in Table 3.1a and Table 3.1b. Only relevant MatInspector matrix matches covering the region of the SNP of interest (at position 124) are included in Table 3.1a and 3.1b.

Table 3.1a Variant Sequence: rs11573156 C

Matrix Family	Description	Start Position	End Position	Anchor	Matrix Similarity	Sequence (Core Sequence in Capitals)
V\$NOLF	Neuron specific olfactory factor	106	128	117	0.884	ccaaccTCCcagagggagcagct
V\$ZNF01	C ₂ H ₂ Zinc finger transcription factors 1	115	139	127	0.781	cagaggggaGCAGctatttaagggga

Table 3.1b Variant Sequence: rs11573156 G

Matrix Family	Description	Start Position	End Position	Anchor	Matrix Similarity	Sequence (Core Sequence in Capitals)
V\$EGRF	EGR/nerve growth factor induced protein C and related factors	113	129	121	0.883	cccagaGGGAggagcta
V\$SP1F	GC-Box factors SP1/GC	114	130	122	0.899	ccagagGGAGgagctat
V\$PLAG	Pleomorphic adenoma gene	119	141	130	0.886	ggGAGGagctatttaaggggagc

Matrix family: describes the transcription factor family predicted by MatInspector, V\$ is the abbreviation for vertebrae and the following three letters are an abbreviation for the family of transcription factor and the fourth letter in the sequence is either 'F'= factor or 'G'= gene. **Description:** A short overview of the potential transcription factor predicted to bind. **Start/End Position:** Where a transcription factor is predicted to begin binding or end binding. **Anchor:** The anchor position is the designated centre point of the matrix sequence. **Matrix Similarity:** Calculated out of a total of 1. A perfect score suggests a total match between a matrix consensus sequence and the input sequence. **Sequence:** Shows the part of the input sequence isolated by the algorithm and predicted to bind a transcription factor.

Table 3.1a-3.1b. A 240bp consensus sequence including the SNP rs11573156 at position 124 was taken from the Ensembl Genome Browser for input to the MatInspector algorithm. Table (a) shows the MatInspector transcription factor prediction algorithm output for the regions including rs11573156 C and Table (b) shows the same output for rs11573156 G.

Conversely the UCSC Genome browser data, which includes a detailed map of predicted transcription factor binding regions within *PLA2G2A*, shows no sign of any TF binding regions within the vicinity of rs11573156 (Figure 3.3), suggesting the likelihood of the predicted Sp1 TF binding site generated by MatInspector truly existing is relatively low. Since the UCSC Genome Browser also shows that rs11573156 is not predicted to be in an open chromatin region we can also reasonably predict that any functionality associated to the SNP is unlikely to be due to differential acetylation or methylation. An additional feature of the UCSC Genome Browser is the identification of splicing regions within a gene. The position of rs11573156 within *PLA2G2A* is confirmed by the UCSC Genome Browser to be near an alternative splice site, 37bp from the untranslated exon 2 boundary (Figure 3.3).

Alternative splicing may be a mechanism through which this SNP acts to affect functionality. Data from the Ensembl Genome Browser supports this idea since one of the two non-coding exons at the 5' end of *PLA2G2A* is commonly spliced out, causing there to be two different protein coding transcripts, which have been previously identified. Evidence for this differential splicing can also be found in the 'supporting evidence' option for each of the protein coding transcripts in Ensembl; ENST00000375111 contains both non-coding exons and ENST00000400520 does not contain non-coding exon 2. Data comparing the alignment of protein and cDNA sequences to the human genome or by manual curation (by the VEGA/Havana group) of human protein, cDNA and EST sequences was used to determine and validate the sequences for each transcript. ENST00000375111 has supporting protein, DNA and EST evidence and ENST00000400520 has supporting protein and cDNA evidence.

Rs11573156 is in the same region as the spliced exon. Should this exon splicing occur differentially between the rs11573156 variants, this could be an explanation for its functionality as the presence or absence of this non-coding exon could contain important transcription regulatory elements or miRNA binding sites that are important in the control of *PLA2G2A* transcription (Figure 3.4).

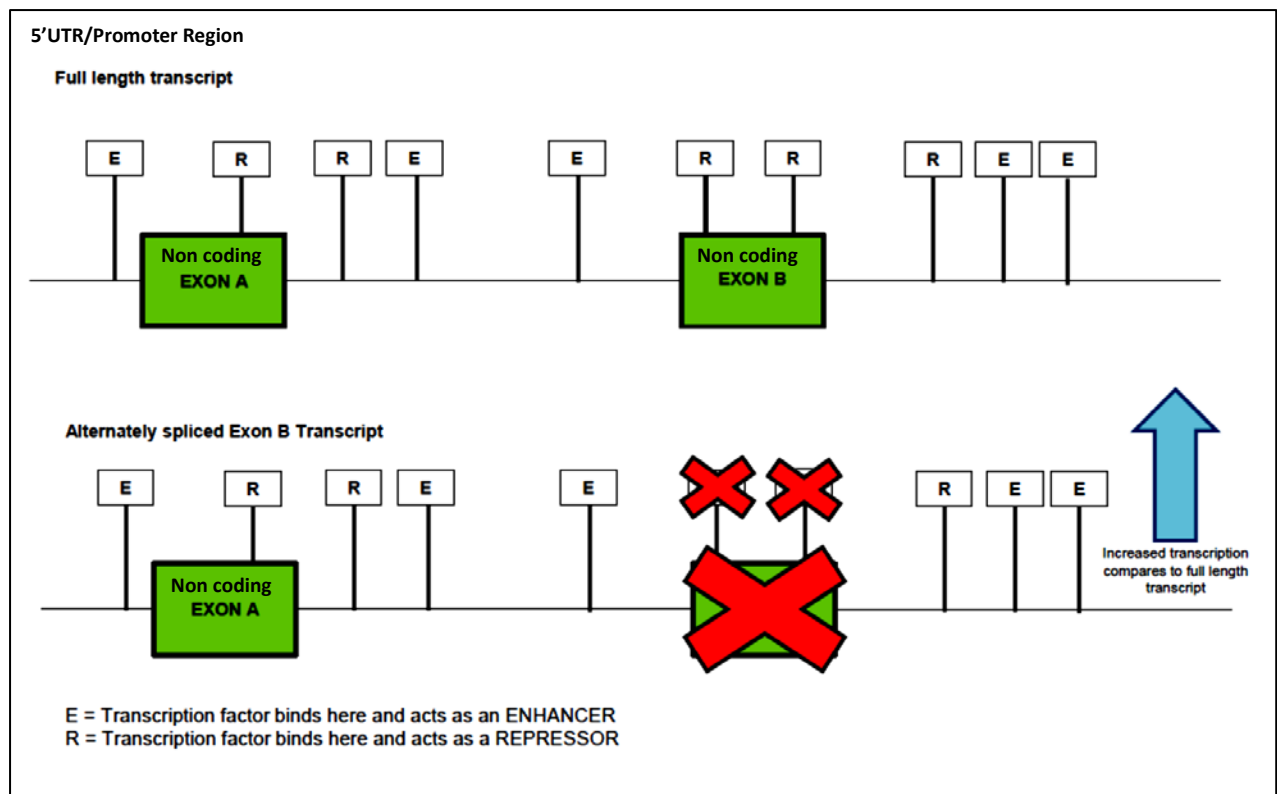


Figure 3.4. Represents an example of alternate splicing. E= Enhancer binding R= Repressor binding. The full length transcript includes 5 transcription factors binding and acting as enhancers and 5 transcription factors binding and acting as repressors. When exon B is alternately spliced out in the second transcript two of the repressor binding sites no longer exist and therefore the transcription factors do not bind here. The lack of two repressor transcription factors increases transcription compared to the full length transcript. If a SNP in this region can be differentially associated with each transcript then this SNP will be marking this alternate splicing.

The chromosomal position of rs11573156 and the surrounding sequence in the untranslated exon 2 has been previously analysed in a deletion mapping study, which employed Chloramphenicol AcetylTransferase (CAT) assay measurements in a human hepatic cell line (HepG2) to compare CAT activity produced by different fragments of the supposed promoter region (Olivier, Fan et al. 1994; Fan, Paradon et al. 1997). The region including rs11573156 and the untranslated exon 2 was determined to contain a negative regulatory region (Olivier, Fan et al. 1994; Paradon, Salvat et al. 1998).

In order to identify potential miRNA binding sites in the 5'UTR of *PLA2G2A*, the miRNA binding site prediction algorithms described in Table 2.4 and **2.4.1** were used. We entered the mRNA sequence for *PLA2G2A*, NM_000300 into the three genome-wide target prediction algorithms that focus outside of the 3'UTR. Using the first algorithm, the Probability of Interaction by Target Accessibility (PITA), we searched for possible microRNA targets (among 470 human microRNAs) that form 8mers with exon 2. There was one

putative miRNA binding site for miR-544 in the region of exon 2, which was in position 96 and had low conservation between species, making it a less likely functional candidate (Table 3.2). In the miRTar algorithm we searched for potential binding sites for 1,104 miRNAs across the selected region. The parameters included information such as minimum free energy of hybridization between microRNAs and their predicted target site (minimum free energy (MFE) of <14Kcal/mol) and an alignment score of ≥ 140 based on structural information on the putative hybrid. There were 13 predicted miRNA binding sites for exon 2 (Table 3.3). The third algorithm, RegRNA, identifies miRNA target sites against an input mRNA sequence, which was the RefSeq mRNA ID NM_00300 in our case. As before, we used as cut-off points MFEs <14Kcal/mol and an alignment score of ≥ 140 which are estimated from the miRanda algorithm. Results included the 13 miRTar predicted miRNA binding sites for exon 2 plus an additional 15 predicted sites (Table 3.4). From the above algorithms, which provided overlapping data to a certain extent, we were able to identify several potential miRNA binding sites. Absence of the active miRNA binding site, in the case of exon skipping, could account for the loss of post-transcriptional repression (Moore, Johnston et al. 2011).

Sites Predictions

Organism	RefSeq	Gene Name	microRNA	Position	Seed	dGduplex	dGopen	ddG	Conservation
Human	NM_000300	PLA2G2A	hsa-miR-765	194	8:0:0	-24	-6.36	-17.63	0
Human	NM_000300	PLA2G2A	hsa-miR-219-1-3p	204	8:0:0	-17.75	-5.5	-12.24	0
Human	NM_000300	PLA2G2A	hsa-miR-9	219	8:0:0	-13.9	-4.97	-8.92	0.00012
Human	NM_000300	PLA2G2A	hsa-miR-544	96	8:0:0	-12	-10.01	-1.98	0
Human	NM_000300	PLA2G2A	hsa-miR-302f	258	8:0:0	-7.8	-9.07	1.27	0.58

Site Predictions; Predict miRNA binding to a single site along the input sequence. **RefSeq;** represents the reference sequence database sequence input for analysis. **Position;** anchor base position of the predicted miRNA binding site relative to the input sequence. **Seed;** the first digit represents the size of the miRNA (8mer), the second digit represents the amount of base mismatches predicted and the third digit represents how many GU wobble base pairs exist. **dGduplex;** ΔG free energy score for the DNA duplex. **dGopen;** ΔG free energy score comparing constrained and non-constrained structures in the sequence. **ddG;** $\Delta\Delta G$ free energy score. The lower this score is the more likely a miRNA is to bind here, As a rule of thumb -10 and below suggests a functional site.

Table 3.2. Represents the output from the PITA miRNA target site prediction algorithm. We searched for 8-mer targets using the RefSeq NM_000300, representing the full length *PLA2G2A* transcript. Exon 2 lay between position 56-166 of the sequence and so only predicted sites between these positions were considered relevant. This left only miR-544 as a potentially relevant miRNA.

miRNA target on Exon #2 in UTR region of the transcript: NM_000300

miRNA	Pos. on Gene	Transcript (Tx.)	Pos. on Tx.	Tx. Region			Exon Type		Conservation	Target Accessibility	MFE	Alignment Score	Alignment	Seed Type
				5'UTR	CDS	3'UTR	CS	AS						
hsa-miR-1976	768 - 787	NM_000300	97 - 116	✓			✓		0.000	0.001	-20.40	164.00	miRNA 3'-ugucUUUCUCC <u>GUCCU</u> C-5' mRNA 5'-uuuuUAGGGGAGCAGGA <u>Gu</u> -3'	7mer-m8
hsa-miR-30b*	736 - 757	NM_000300	65 - 86	✓			✓		0.000	0.038	-20.90	151.00	miRNA 3'-caucauuUGUAGG-UGGAGGGU <u>C</u> -5' mRNA 5'---uacacuACCCCAAC <u>UCCCA</u> g-3'	7mer-m8
hsa-miR-3192	737 - 758	NM_000300	66 - 87	✓			✓		0.000	0.026	-23.80	166.00	miRNA 3'-aagGUGACCAU---GUUCCAGCC <u>GU</u> C-5' mRNA 5'---CAC--CUACCCCAAC <u>UCCCA</u> g-3'	8mer
hsa-miR-326	741 - 760	NM_000300	70 - 89	✓			✓		0.000	0.003	-16.20	141.00	miRNA 3'-gacuuuuUUCCCG <u>GUCCU</u> C-5' mRNA 5'-uacuuuuAACCUCCAGAG <u>g</u> -3'	7mer-m8
hsa-miR-330-5p	739 - 760	NM_000300	68 - 89	✓			✓		0.001	0.003	-16.20	140.00	miRNA 3'-cggaucuguguccGGGUG <u>GU</u> C-5' mRNA 5'-cuuacuuuuuuCCAGAG <u>g</u> -3'	7mer-m8
hsa-miR-377*	733 - 754	NM_000300	62 - 83	✓			✓		0.000	0.057	-15.30	150.00	miRNA 3'-cuuaAGUGG-UUCCCGUCCAG <u>A</u> -5' mRNA 5'---gcUACCUAACCCCAAC <u>UCC</u> -3'	7mer-m8
hsa-miR-4268	835 - 1555	NM_000300	164 - 184	✓			✓		0.009	0.000	-18.60	150.00	miRNA 3'-gagagggacucUCCUCCUG <u>G</u> -5' mRNA 5'-gagagggacacAGGAGGAG <u>C</u> -3'	8mer
hsa-miR-4314	742 - 759	NM_000300	71 - 88	✓			✓		0.000	0.015	-19.60	146.00	miRNA 3'-gacaGGGUAAAGGGUCC <u>U</u> -5' mRNA 5'-aacccCAACCUCCAGAG <u>g</u> -3'	7mer-m8
hsa-miR-483-3p	768 - 789	NM_000300	97 - 118	✓			✓		0.000	0.001	-20.60	158.00	miRNA 3'-uacagccUCCUC-UCCUAC <u>C</u> -5' mRNA 5'-uuuuuuuAGGGAGCAGGAG <u>U</u> C-3'	7mer-m8
hsa-miR-508-5p	807 - 828	NM_000300	136 - 157	✓			✓		0.000	0.000	-18.80	160.00	miRNA 3'-gaACUACUGCG-GGAGACCUCA <u>u</u> -5' mRNA 5'---cUGGC-CAUACAACUCUCCAG <u>U</u> -3'	7mer-m8
hsa-miR-520a-5p	806 - 826	NM_000300	135 - 155	✓			✓		0.000	0.000	-17.20	140.00	miRNA 3'-ucuuuauagaaggGAGACCU <u>C</u> -5' mRNA 5'-ccuggggauacaaCUCCGGA <u>g</u> -3'	7mer-m8
hsa-miR-525-5p	805 - 826	NM_000300	134 - 155	✓			✓		0.000	0.000	-15.20	146.00	miRNA 3'-ucuuuCAACUAGG--GAGACCU <u>C</u> -5' mRNA 5'---gcuuGGGGAUACAACUCUGGA <u>g</u> -3'	7mer-m8
hsa-miR-92a-1*	731 - 752	NM_000300	60 - 81	✓			✓		0.001	0.057	-16.20	148.00	miRNA 3'-ucgaaACG--UUGGCU-AGGGUUGGA <u>-5'</u> mRNA 5'-----cUCCUACCUACCCCAAC <u>C</u> -3'	7mer-m8

Exon Type: AS = alternatively spliced exon CS = constitutively spliced. **Target Accessibility;** Predicted miRNA binding sites in more accessible parts of the DNA sequence are more likely to be accurate. **Alignment Score;** Score calculated by the MiRanda algorithm- a weighted score of the match and mismatch base pairs in the sequence with gap penalties (Higher = more likely to be truly binding).

Table 3.3. Represents the output from the miRTar miRNA target site prediction algorithm. We searched for 8-mer targets using the RefSeq NM_000300, representing the full length *PLA2G2A* transcript. Parameters of minimum free energy ≤ -14 Kcal/mol and a Score of ≥ 140 were used.

miRNA ID	Location	Hybridization	MFE	Score
hsa-miR-1224-5p	144-162	miRNA: 3' ggUGGAGGGCUCAGGAGug 5' : Target:5' acAACTCTGGAGTCCTctg 3'	-21.2	145
hsa-miR-1236	159-183	miRNA: 3' gacCUCUCU-----GUUCCCCUUCUCc 5' : Target:5' totGAGAGAGCCACCA--GGAGGAGc 3'	-19.9	143
hsa-miR-18a*	83-110	miRNA: 3' ggUCUCCUCG---UGAA-UCCCGUCa 5' : : Target:5' ccAGAGGGAGCAGCTATTTAAGGGGAGc 3'	-28	143
hsa-miR-1976	97-116	miRNA: 3' ugucGUUCCUCCCGUCCUCc 5' : : Target:5' tattTAAGGGGAGCAGGAGt 3'	-20.4	164
hsa-miR-1976	169-189	miRNA: 3' ugUCGUUCCUCC-CGUCCUCc 5' : Target:5' ccACCAAGGAGGAGCAGGGGa 3'	-27.7	158
hsa-miR-19b-1*	102-124	miRNA: 3' cgaCCUACGUUUGGACGUUUUGa 5' : : : : Target:5' aagGGGAGCAGGAGTGCAGAACa 3'	-17.8	148
hsa-miR-19b-2*	103-124	miRNA: 3' acuuuaCGUUUGGACGUUUUGa 5' : : : Target:5' aggggaGCAGGAGTGCAGAACa 3'	-15.6	140
hsa-miR-26b*	106-126	miRNA: 3' cucgGUUCAUUAACCUCUUGUcc 5' : : Target:5' ggagCAGG-AGTGCAGAACAaa 3'	-14.4	141
hsa-miR-30b*	64-86	miRNA: 3' cuucauuUGUAGG-UGGAGGGUc 5' Target:5' ctcacctACCCCCAACCTCCCAg 3'	-20.9	151
hsa-miR-3183	150-170	miRNA: 3' aggCUCGCUGAGGCUCUCUCcg 5' : : Target:5' ctgGAGT-CCTCTGAGAGAGcc 3'	-22.3	146
hsa-miR-3192	61-87	miRNA: 3' aaGGUGACGAU----GUUGGAGGGUCu 5' Target:5' tgCCTCACCTACCCCCAACCTCCCAGa 3'	-24.1	167
hsa-miR-324-3p	187-206	miRNA: 3' ggUCGUCGUGGACCCCGUCa 5' : : Target:5' ggAGCGACGCCGGGGCAGa 3'	-21.8	158
hsa-miR-326	70-89	miRNA: 3' gaccuccUUCCCGGGUCUCc 5' Target:5' taccocccAACCTCCCAGAGg 3'	-16.2	141
hsa-miR-330-5p	68-89	miRNA: 3' cggauucuguguccGGGUCUCu 5' Target:5' cctacccccaaacctCCCAGAGg 3'	-16.2	140
hsa-miR-377*	61-83	miRNA: 3' cuuaAGUGG-UUCCCGUUGGAGa 5' Target:5' tgccTCACCTACCCCCAACCTCc 3'	-15.3	150

hsa-miR-4268	164-184	miRNA: 3' guguaggacucUCCUCCUCGg 5' Target:5' gagagccaccaAGGAGGAGCa 3'	-18.6	150
hsa-miR-431*	117-138	miRNA: 3' ucuucgggacGUUCUGCUGGAc 5' Target:5' gcagaacaaaCAAGACGGCCTg 3'	-22.3	144
hsa-miR-4314	71-88	miRNA: 3' gacaGGGUAAGGGUCUc 5' Target:5' acccCCAACCTCCCAGAg 3'	-19.6	146
hsa-miR-4326	158-184	miRNA: 3' caGACCCUCU-----GU--CUCCUUGu 5' Target:5' ctCTGAGAGAGCCACCAAGGAGGAGCa 3'	-18.3	140
hsa-miR-483-3p	97-118	miRNA: 3' uucugccCUCCUC-UCCUCACu 5' Target:5' tatattaaGGGGAGCAGGAGTGc 3'	-20.6	158
hsa-miR-508-5p	135-157	miRNA: 3' guACUCACUGCG--GGAGACCUCAu 5' Target:5' ccTGGG-GATACAACTCTGGAGTc 3'	-18.8	160
hsa-miR-516a-3p	77-94	miRNA: 3' ugGGAGACUUCCUUCGu 5' Target:5' aaCCTCCCAGAGGGAGCa 3'	-20.1	144
hsa-miR-516b*	77-94	miRNA: 3' ugGGAGACUUCCUUCGu 5' Target:5' aaCCTCCCAGAGGGAGCa 3'	-20.1	144
hsa-miR-520a-5p	135-155	miRNA: 3' uuuuugaaggGAGACCUc 5' Target:5' cctgggatacaaaCTCTGGAg 3'	-17.2	140
hsa-miR-525-5p	133-155	miRNA: 3' uuuuacacguagg--GAGACCUc 5' Target:5' ggcctGGGATACAACTCTGGAg 3'	-15.2	146
hsa-miR-562	83-102	miRNA: 3' cgUUUACCAUGUCGAUGAA 5' Target:5' ccAGAGGGAGCAGCTATTTa 3'	15.3	150
hsa-miR-637	64-87	miRNA: 3' ugcgucucGGGCUUUCGGGGGUCa 5' Target:5' ctcacctaCCCCAACCTCCCAGa 3'	-16.7	140
hsa-miR-92a-1*	54-81	miRNA: 3' ucGUAACGU-----UGGCU-AGGGUUGGa 5' Target:5' agCCTTGTGCCTCACCTACCCCCAACCT 3'	-17.4	151
hsa-miR-938	85-108	miRNA: 3' ugaCCCAAGUGG--AAAUUCCGgu 5' Target:5' agaGGGAGCAGCTATTTAAGGGga 3'	-17.4	141

Table 3.4. Output from RegRNA algorithm for *PLA2G2A* NM_000300. Results for the region of exon 2 are shown in the table. Parameters MFE ≤ -14Kcal and Score ≥ 140. Prediction Score is based on the MiRanda algorithm. Hybridization = sequence predicted to bind a miRNA. MFE= minimum free energy. A lower MFE means an increased chance of binding. The score is weighted on matched and mismatched pairs and includes gap penalties. Scores above 140 are considered more likely to represent true binding sites.

In summary, 11 miRNAs are predicted to bind to the exon 2 region of *PLA2G2A* by two out of the 3 algorithms used. These are; hsa-mir-1976, hsa-mir-3192, hsa-mir-326, hsa-mir-330-5p, hsa-mir-4268, hsa-mir-4314, hsa-mir-483-3p, hsa-mir-508-5p, hsa-mir-520a-p, hsa-mir-525-5p, hsa-mir-92a-1. Of these, two miRNAs are predicted to bind in the region of rs11573156; hsa-mir-1976 and hsa-483-3p. If rs11573156 genotype is associated with the alternate splicing of exon 2 then any miRNA binding here will also be associated with exon 2 alternate splicing, since the splicing of exon 2 (and rs11573156) would result in the removal of the miRNA target seed region.

3.2.3. Bioinformatics Study of rs3767221

Data from NCBI and Ensembl verifies that rs3767221 is positioned at chromosome 1:20,301,781, which is in the 3'UTR of *PLA2G2A*. The European genotyping data available from the 1000 Genomes Project showed that the common allele, A, occurs at a frequency of 0.59 and the rare C allele occurs at a frequency of 0.41. The linkage disequilibrium between rs3767221 and rs11573156 is quoted as $R^2 = 0.076$ in the 1000 Genomes CEU pilot 1 data, which is not significant. However, since rs3767221 is a tagging-SNP it is in strong LD with several other SNPs. Those with an R^2 above 0.7 are rs10799599, rs2872821, rs2315070, rs12023742 and rs12044628 (all of these SNPs lie between *PLA2G2A* and *PLA2G2E*) and rs1890686 (*PLA2G2A* 3'UTR). There is no European population HapMap Project data for rs3767221 and therefore the Haploview 4.0 heatplot tool cannot be used to generate an LD heatplot. Figure 3.5 shows an LD Manhattan style plot generated using the SNAP Pairwise LD Plot tool (2.2 Table 2.4), using the following parameters; Regional LD Plot for plot type and Text Entry for SNP input; rs3767221 was the input SNP. The R^2 value was set to 0.7 and the 1000 Genomes Pilot 1 CEU dataset was selected. Finally the distance limit input was set to 100Kb.

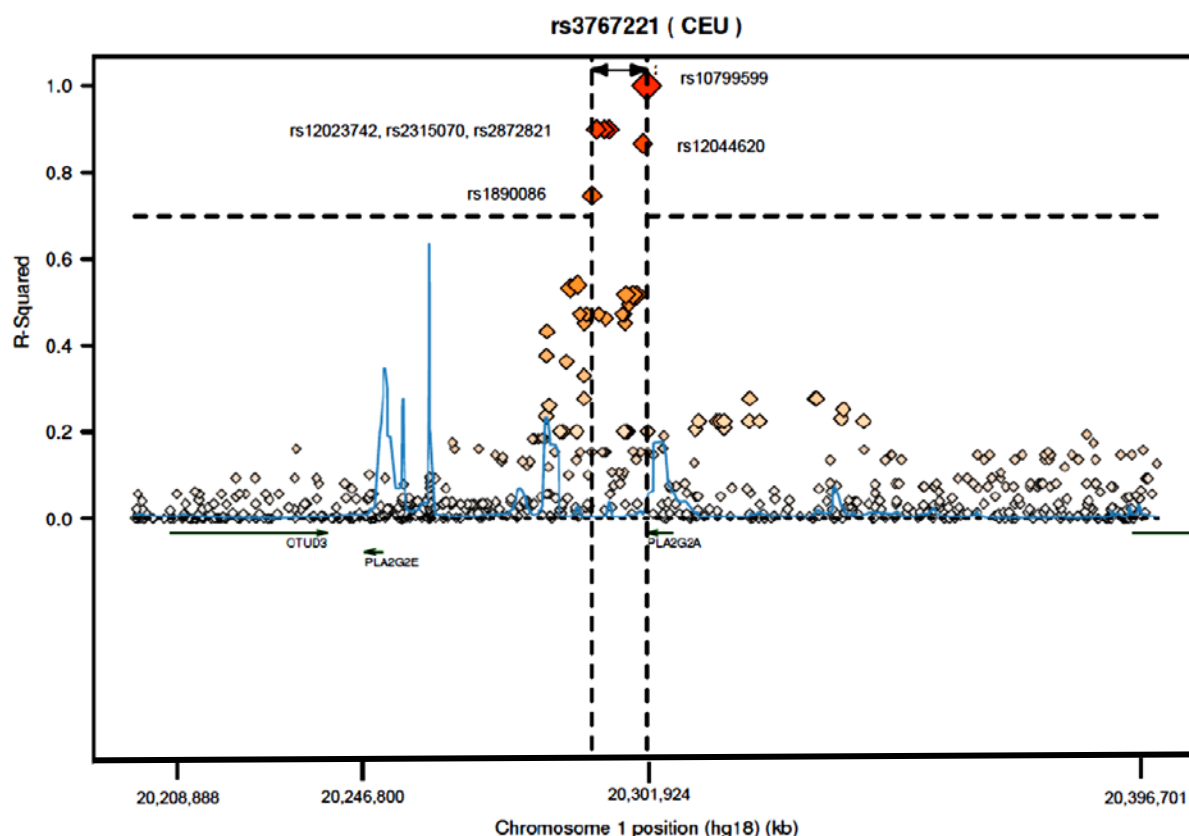


Figure 3.5. An LD Manhattan style plot based on the 1000 Genomes CEU pilot 1 data available through the SNAP Pairwise LD program. The SNAP plot tool was used to assess the LD of rs3767221 with all known SNPs in the region of 100Kb up and downstream of the SNP. The region encompassing the SNPs in LD <0.7 with rs11573156 is represented by the two vertical dashed lines. Genotyped and imputed SNPs are represented by a diamond shape. The R^2 significance value is set at 0.7 and represented by a horizontal dashed line. The SNPs in LD above 0.7 are indicated above this line. The deeper the shade of red and the larger the size of each diamond, the higher the strength of LD with rs3767221. The blue peaks on the plot represents recombination hotspots and the green arrows represent different genes.

Prediction by MatInspector was carried out using a 240 base sequence taken from Ensembl Genome Browser including rs3767221 on the reverse strand. Table 3.5a and 3.5b show the results from this MatInspector algorithm output that cover the region including rs3767221 (position 172). This prediction algorithm suggests the strongest potential differential binding to rs3767221 C compared to A is a potential cAMP binding protein site with a matrix similarity of over 0.85 in the region of rs3767221 C that is not predicted for rs3767221 A. However the anchor base for this binding site is not within 5bp of rs3767221. The only prediction with a matrix similarity of over 0.9 and within 5 bases of rs3767221 was suggested to be a growth factor independence transcriptional repressor binding to rs3767221 A at its anchor position, but not at rs3767221 C.

Table 3.5a rs3767221 A

Matrix Family	Description	Start Position	End Position	Anchor	Matrix Similarity	Sequence (Core Sequence in Capitals)
V\$GFI1	Growth factor independence transcriptional repressor	165	179	172	0.913	agcAATCattgcact
V\$CAAT	CCAAT binding factors	168	182	175	0.812	tcaaGCAAtcattgc
V\$HOXC	HOX-PBX complexes	168	184	176	0.807	gcaatGATTgcttgagc

Table 3.5b rs3767221 C

Matrix Family	Description	Start Position	End Position	Anchor	Matrix Similarity	Sequence (Core Sequence in Capitals)
V\$CREB	cAMP-responsive element binding proteins	156	176	166	0.876	tgaggcTGAAgtgcaaggatt
V\$AHRR	AHR-arnt heterodimers and AHR-related factors	167	191	179	0.775	tgcaaggattGCTTgagctcacaga

Matrix family: describes the transcription factor family predicted by MatInspector, V\$ is the abbreviation for vertebrae and the following three letters are an abbreviation for the family of transcription factor and the fourth letter in the sequence is either 'F'= factor or 'G'= gene. **Description:** A short overview of the potential transcription factor predicted to bind. **Start/End Position:** Where a transcription factor is predicted to begin binding or end binding. **Anchor:** The anchor position is the designated centre point of the matrix sequence.

Matrix Similarity: Calculated out of a total of 1. A perfect score suggests a total match between a matrix consensus sequence and the input sequence. **Sequence:** Shows the part of the input sequence isolated by the algorithm and predicted to bind a transcription factor.

Table 3.5a-3.5b. A 240bp consensus sequence including the SNP rs3767221 at position 172 was taken from the Ensembl Genome Browser for input to the MatInspector algorithm. Table (a) shows the MatInspector transcription factor prediction algorithm output for the regions including rs3767221 A and Table (b) shows the same output for rs3767221 C.

The UCSC Genome Browser predicts a dense protein binding region in the sequence around rs3767221 (Figure 3.6), by *in vitro* assays and computer based algorithms. This region included predicted binding sites for SREBP 1 and sPLA2 interacting transcription factors such as p300 (Giandomenico, Simonsson et al. 2003; Ponugoti, Kim et al. 2010) and LXR alpha (Trousson, Bernard et al. 2009). The UCSC Genome Browser also reports several other known transcription factor binding regions at this site including; RXRA, CEBPB, HNF4A and p300, and the Aryl hydrocarbon receptor (AHR). Since transcription factors tend to bind to the 5'UTR it is likely that predicted binding sites are in fact for proteins other than transcription factors.

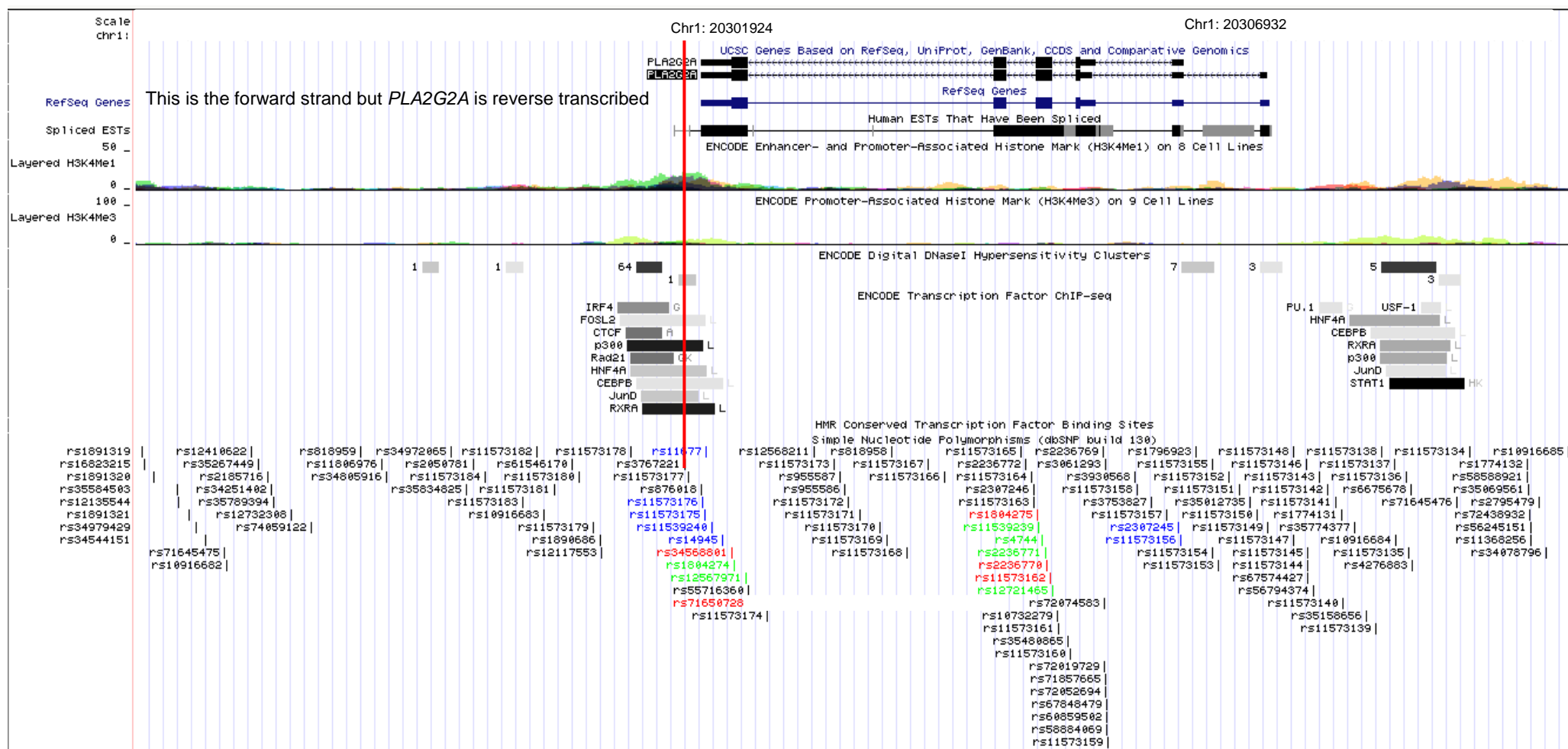


Figure 3.6. Has been reprinted using data from the UCSC Genome Browser for *PLA2G2A*. The gene is shown on the forward strand but is usually reverse transcribed. The full transcript is shown based on data from RefSeq, UniProt, GenBank, CCDS and Comparative Genomics. Spliced ESTs represents Human expressed sequence tags (short strands of cDNA used to identify gene transcripts), which have been used to identify potential splice sites within the gene. These are shown in black. Layered H3K4Me1 represents data from Encode ChIP sequence assays to show the level of predicted histone 3 and 4 methylation/acetylation. Each colour represents a different cell type in which the assay was carried out: Dominant in this schematic, are human mammary epithelial cells (HMEC) (green), and human embryonic stem cell line (H1 ES) (yellow). Layered H3K4Me3 shows DNase hypersensitivity regions are shown in the same cell lines. The ENCODE Transcription Factor ChIP-seq data shows the prediction of transcription factors binding to the gene. These are labelled accordingly. The intensity of colour indicated the likelihood of true binding. The cells lines represented here are: G/GK=GM12872, A= A549, L= HepG2, HK= HeLa-S3. Finally all the SNPs 162 covering the region are indicated by rs number at the bottom. The SNP of interest, rs3767221, is indicated by the vertical red line (<http://genome.ucsc.edu/cgi-bin/hgTracks?position=chr1:20174118-20174618&hgid=284369053&snp130=pack&hgFind.matches=rs3767221>).

Micro-RNA binding usually occurs in the 3'UTR. The SNP rs3767221 is situated in the 3'UTR of *PLA2G2A*. Should a miRNA bind differentially to the region of rs3767221 this SNP could be a post-transcriptional regulator of gene expression. A tentative potential miRNA binding site in the region of rs3767221 is hsa-miR-1273g-3p. Interestingly, within the recognised binding site for this miRNA; ACCACUGCACUCCAGCCUGAG is a sequence similar to the recognised SREBP binding sequence; TCACNCCAC (base pair matches highlighted in yellow). This suggests that miRNAs with similar recognised binding sequences to transcription factors could be truly binding in the region of UCSC Genome browser predicted transcription factor binding in the 3'UTR. Table 3.6 shows the results from miRBase. Results returned 935 predicted miRNA binding sites within the 801bp region, 22 of which were within 5 bases up or downstream of our SNP of interest (rs3767221 was at position 401) and are shown below in Table 3.7.

Accession	ID	Query start	Query end	Subject start	Subject end	Score	Evalue
MIMAT0005876	hsa-miR-1285-3p	67	87	2	22	87	0.04
MIMAT0022719	hsa-miR-1285-5p	68	88	1	21	78	0.22
MIMAT0022286	hsa-miR-5585-3p	143	163	1	21	69	1.3
MIMAT0015001	hsa-miR-3135a	125	141	6	22	67	1.8
MIMAT0015090	hsa-miR-1273d	126	142	9	25	67	1.8
MIMAT0019232	hsa-miR-4423-5p	211	227	6	22	67	1.8
MIMAT0022285	hsa-miR-5585-5p	140	161	1	22	65	2.7
MIMAT0022742	hsa-miR-1273g-3p	118	138	1	21	65	2.7
MIMAT0020602	hsa-miR-1273g-5p	126	143	1	18	63	4
MIMAT0022708	hsa-miR-584-3p	84	101	3	20	63	4
MIMAT0005926	hsa-miR-1273a	61	85	1	25	62	4.8
MIMAT0004952	hsa-miR-665	132	145	3	16	61	5.8
MIMAT0019049	hsa-miR-4512	66	86	2	22	60	7.1
MIMAT0019216	hsa-miR-3187-5p	127	147	2	22	60	7.1
MIMAT0019697	hsa-miR-4639-5p	129	149	2	22	60	7.1
MIMAT0022473	hsa-miR-5684	78	96	1	19	59	8.6
MIMAT0025458	hsa-miR-6501-5p	79	97	2	20	59	8.6

Table 3.6 Results for miRBase analysis of a 240bp region the quoted NCBI sequence for rs3767221 with the SNP located at position 121. The sequence was blasted using the parameters: Mature miRNA for search sequences option, search method BLASTN, E-Value cut off set at 10 (this value represents the match efficiency between the input sequence and the predicted miRNA sequence, the lower the score the higher the match efficiency so an e value of 0 = 100% match), maximum number of hits set at 100 and human specific results were selected for. The accessions in the table are for the miRBase database details of each miRNA. Query start and end relate to the input sequence position matched to a miRNA sequence. Subject start and end refers to the base position of the miRNA sequence matched to the input sequence. The score relates to the miRBase algorithm calculation weighting matches to mismatches and taking into account gap penalties. The higher the score the more likely a prediction is a true miRNA binding sequence. The e value represents the number of mismatches between the two sequences. A lower score is better with 0 representing a 100% match between sequences.

Gene	microRNA	Position	Seed	dGduplex	dGopen	ddG
PLA2G2A	hsa-miR-944	396	8:1:0	-5.5	-7.06	1.56
PLA2G2A	hsa-miR-522	397	8:1:1	-8.2	-7.06	-1.13
PLA2G2A	hsa-miR-518a-5p	399	8:1:1	-8.81	-7.06	-1.74
PLA2G2A	hsa-miR-527	399	8:1:1	-8.81	-7.06	-1.74
PLA2G2A	hsa-miR-573	399	8:1:1	-8.3	-7.06	-1.23
PLA2G2A	hsa-miR-130b	401	8:0:0	-21.7	-8.64	-13.05
PLA2G2A	hsa-miR-454	401	8:0:0	-19.8	-8.64	-11.15
PLA2G2A	hsa-miR-301b	401	8:0:0	-18.4	-8.64	-9.75
PLA2G2A	hsa-miR-130a	401	8:0:0	-18.12	-8.64	-9.47
PLA2G2A	hsa-miR-512-3p	401	8:1:1	-15.9	-8.64	-7.25
PLA2G2A	hsa-miR-301a	401	8:0:0	-15.5	-8.64	-6.85
PLA2G2A	hsa-miR-372	402	8:1:1	-17.3	-9.66	-7.63
PLA2G2A	hsa-miR-519b-3p	402	8:1:0	-15.56	-9.66	-5.89
PLA2G2A	hsa-miR-519c-3p	402	8:1:0	-13.24	-9.66	-3.57
PLA2G2A	hsa-miR-519a	402	8:1:0	-11.97	-9.66	-2.3
PLA2G2A	hsa-miR-423-5p	403	8:1:1	-14.3	-9.74	-4.55
PLA2G2A	hsa-miR-18b	403	8:1:1	-13.8	-9.74	-4.05
PLA2G2A	hsa-miR-18a	403	8:1:1	-13.6	-9.74	-3.85
PLA2G2A	hsa-miR-1289	403	8:1:1	-10.8	-9.74	-1.05
PLA2G2A	hsa-miR-573	404	8:1:0	-12.1	-10.67	-1.42
PLA2G2A	hsa-miR-203	404	8:1:1	-9.41	-10.67	1.26
PLA2G2A	hsa-miR-138	406	8:1:1	-16.4	-10.76	-5.63

Site Predictions; Predict miRNA binding to a single site along the input sequence. **Gene;** represents the gene for which the 3'UTR sequence has been selected for input for analysis. **Position;** anchor base position of the predicted miRNA binding site relative to the input sequence: NM_000300. **Seed;** the first digit represents the size of the miRNA (8mer), the second digit represents the amount of base mismatches predicted and the third digit represents how many GU wobble base pairs exist. **dGduplex;** ΔG free energy score for the DNA duplex. **dGopen;** ΔG free energy score comparing constrained and non-constrained structures in the sequence. **DdG;** $\Delta\Delta G$ free energy score. The lower this score is the more likely a miRNA is to bind here, as a rule of thumb for the PITA algorithm -10 and below suggests a functional site (http://genie.weizmann.ac.il/pubs/mir07/mir07_notes.html).

Table 3.7. Represents the output from the PITA miRNA target site prediction algorithm. We searched for 8-mer targets using the Ensembl SNP tool for the rs3767221 flanking sequence, representing part of the 3'UTR *PLA2G2A* transcript. The SNP rs3767221 lay at position 401 of the sequence and so only predicted sites at this position or 5 bases up or downstream were considered relevant.

Bioinformatics tools were also used to check the other previously identified *PLA2G2A* tSNPs to examine if they might influence functional domains, but these results were negative and we did not continue to investigate them further.

3.2.4. Functional Experiments

3.2.4.1. Luciferase Assay Results for Analysis of rs11573156

The luciferase constructs for rs11573156 C and G were isolated, ligated into a pGL3-Basic vector and verified as previously described (2.4.7-2.4.7.5). The luciferase assays for rs11573156 C and G alleles showed very low expression, being in the order of 1.2 to 1.6 relative luciferase units, compared to the empty vector alone, with the G allele showing a modest 10% higher level ($P=0.02$) (Figure 3.7).

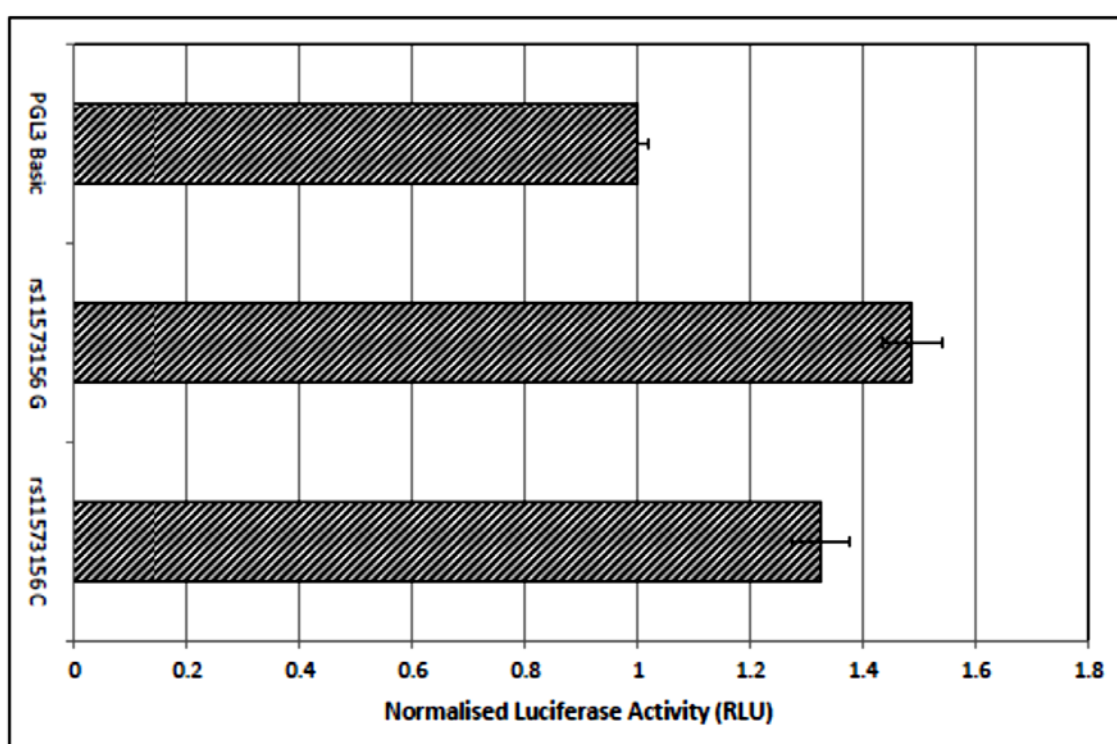


Figure 3.7. The difference in luciferase activity (Relative Light Units) between the C and G alleles of rs11573156 normalised to pGL3-Basic. There is no pre-existing promoter in this vector and both inserted constructs for C and G variants give very low levels of luciferase compared to the baseline measurement of the vector alone. The error bars represent the standard error mean.

3.2.4.2. Luciferase Assay Results for Analysis of rs3767221

Results from the luciferase assays for rs3767221 A>C are shown in Figure 3.8. A significant difference was seen, with the C allele showing ~55% lower luciferase activity compared to the T allele ($P<0.0001$). Luciferase activity measures for these variants were in the order of 27-62 RLU.

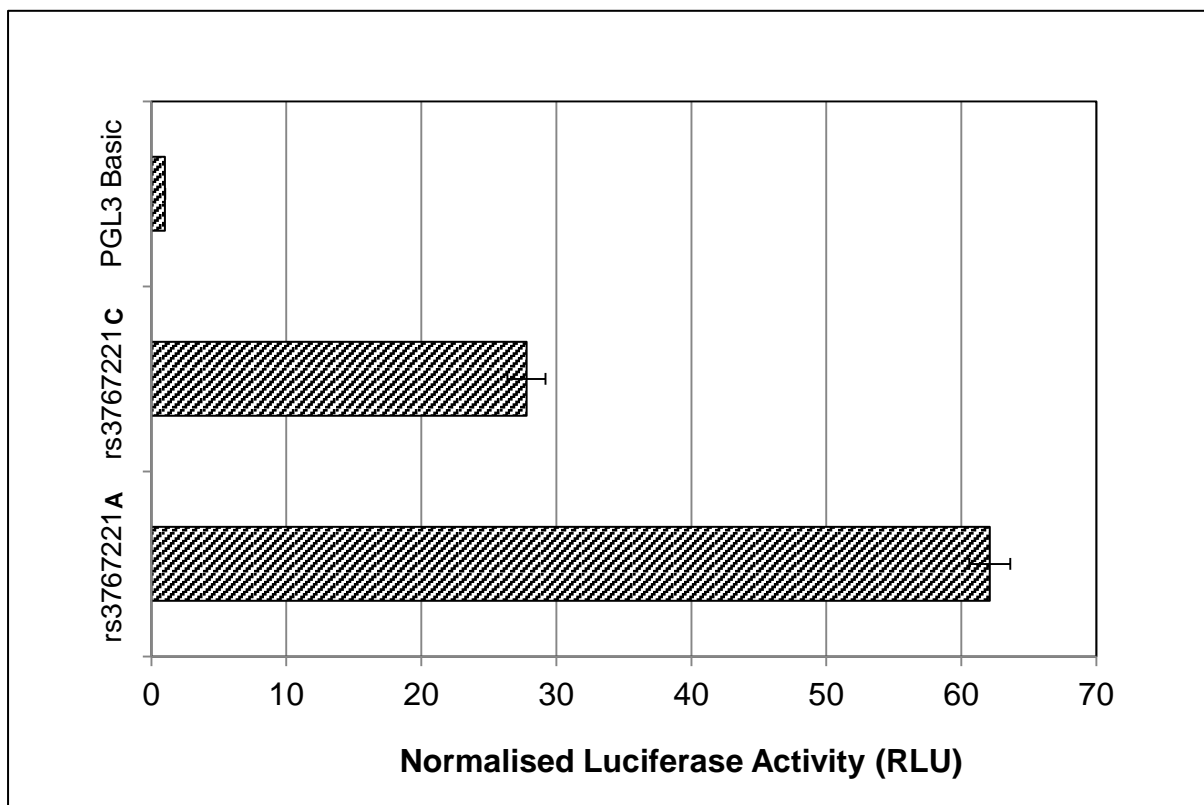


Figure 3.8. The difference in luciferase activity (Relative Light Units) between the A and C alleles of rs3767221 (forward orientation), normalised to pGL3 expression vector. Rs3767221 C (rare) shows a 55% lower Luc activity compared to the A allele (wild type). The empty pGL3-Basic vector is used as a promoter free baseline measurement for transcription. The error bars represent the standard error mean.

3.2.5. EMSA Results

3.2.5.1. EMSA Results for the Analysis of rs11573156

Figure 3.9. Shows a standard EMSA for rs11573156. EMSAs specific to either the rs11573156 C or G alleles did not show any differential, specific nuclear protein binding. The transcription factor binding site prediction algorithm MatInspector identified only Sp1 as a putative differential transcription factor binding sites in the 60bp region flanking rs11573156. Sp1 protein was therefore used as a potential positive control in rs11573156 EMSA analysis, but results showed that a specific Sp1 band was not present for either biotin labelled rs11573156 allele. Examination of this region on the UCSC Genome Browser also showed an absence of predicted transcription factor binding sites (Figure 3.3).

In each EMSA assay sterol regulatory element-binding protein 1 (SREBP 1) was used as a positive control as this had been previously optimised for use in an EMSA by

members of our lab. I could therefore be satisfied that this protein binding pattern could be used to confirm that the EMSA had worked in each instance.

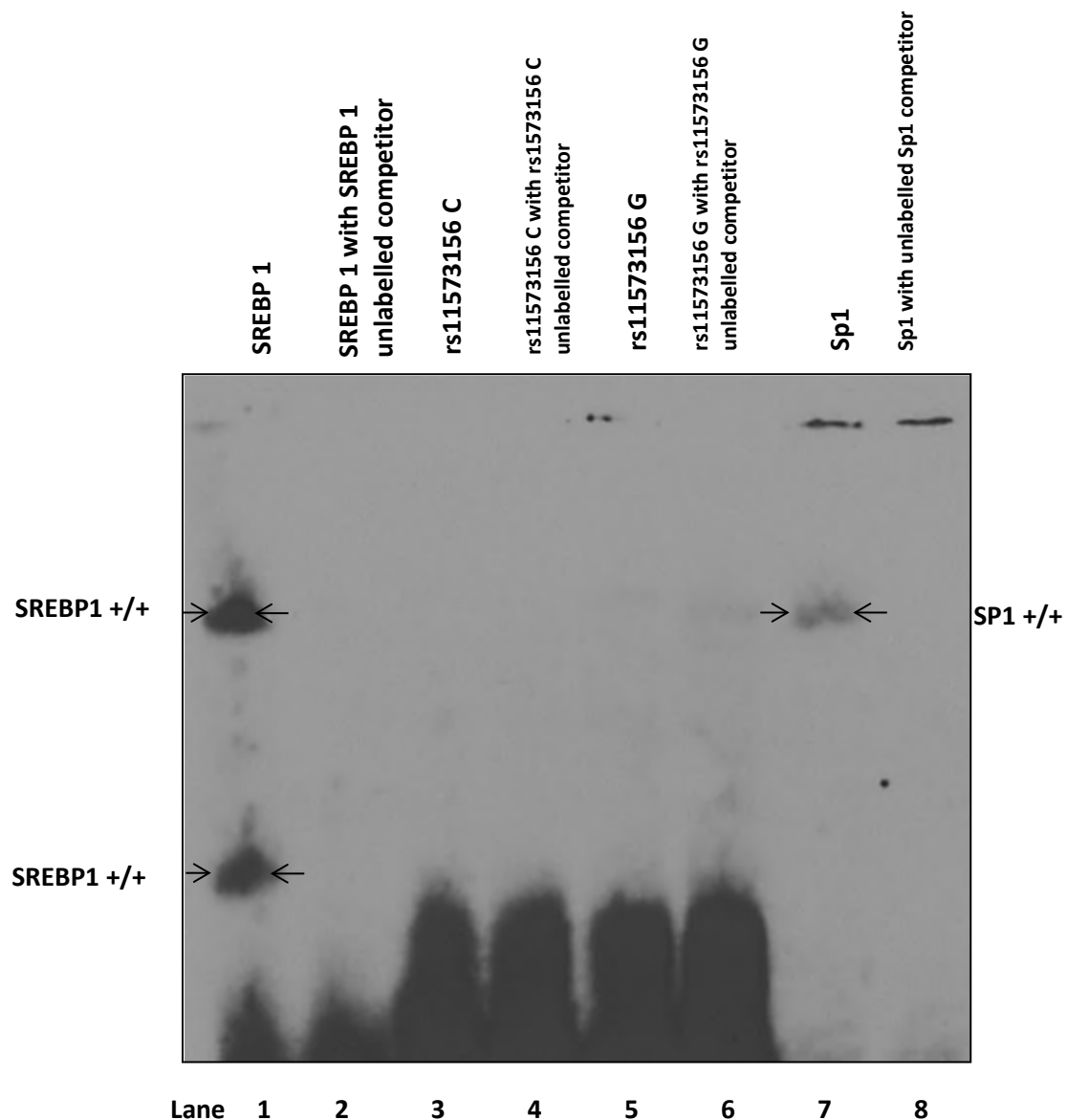


Figure 3.9. EMSA image of rs11573156 C and G alleles showing no transcription factor binding. Lane 1 is SREBP 1 positive control that shows specific bands that are competed out with an unlabelled SREBP 1 consensus sequence in lane 2. Lanes 3 and 4 show rs11573156 C (wild type) labelled and with unlabelled C competitor respectively. There is no specific binding for either. Lanes 5 and 6 represent rs11573156 G (rare) labelled and with unlabelled G competitor respectively, again with no specific binding. Lane 7 represents a second positive control with Sp1 specific binding and in Lane 8, labelled Sp1 with unlabelled Sp1 competitor that competes out the specific bands.

3.2.5.2. EMSA Results for the Analysis of rs3767221

An EMSA (Figure 3.10) confirmed differential binding of a nuclear protein between the A and C alleles, illustrated by the additional band seen with the biotin-labelled rs3767221 C, compared to rs3767221 A. This additional band seen with the C variant was competed out by the un-biotinylated C probe, suggesting specific binding of a protein in the presence of the C allele but not the A allele. The difference in transcription factor binding associated with this 3'UTR SNP could influence the rate of transcription and/or mRNA stability, thus affecting *PLA2G2A* expression and levels of sPLA2-IIA.

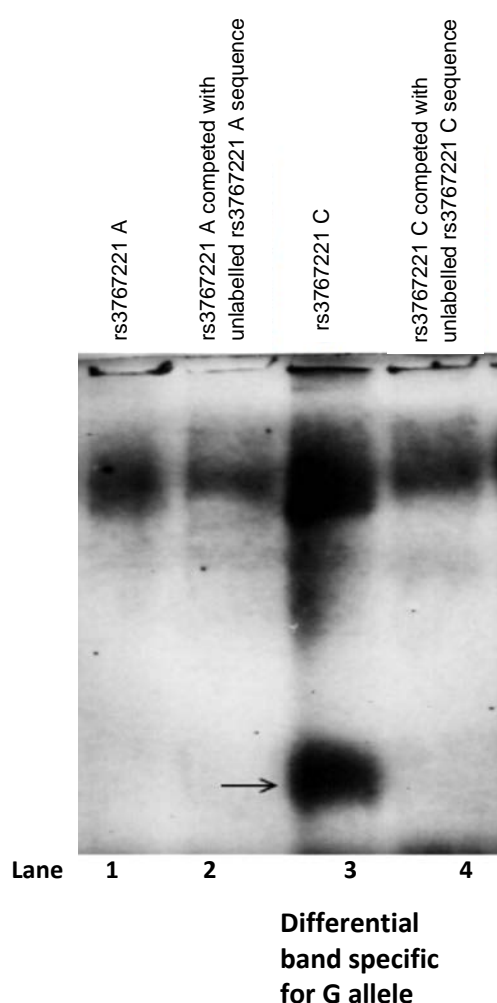


Figure 3.10. Rs3767221 EMSA image showing differential binding between A and C alleles. Lane 1 shows rs3767221 A. Lane 2 shows rs3767221 A competed with unlabelled rs3767221 A. Lane 3 shows rs3767221 C. Lane 4 shows rs3767221 C competed with unlabelled rs3767221 C. A specific band is seen with the rare (C) variant in Lane 3 compared to the wild type (A) variant in Lane 1. In Lane 4, although the competitor reduces the intensity of all bands, only the lower band seen with the A allele is competed out completely.

Coincidentally, it was noted that the differentially binding band was a similar size to one of the SREBP1 control bands. To investigate this further, rs3767221 C samples were competed with SREBP 1. For rs3767221 C, a distinct reduction in the intensity of the band was seen, suggesting that an SREBP 1 related protein may be involved in the differential binding to the alleles of rs3767221 (Figure 3.11). Though it is unlikely to be SREBP1 specifically since the known SREBP1 binding sequence does not appear to be present in the 3'UTR of *PLA2G2A*.

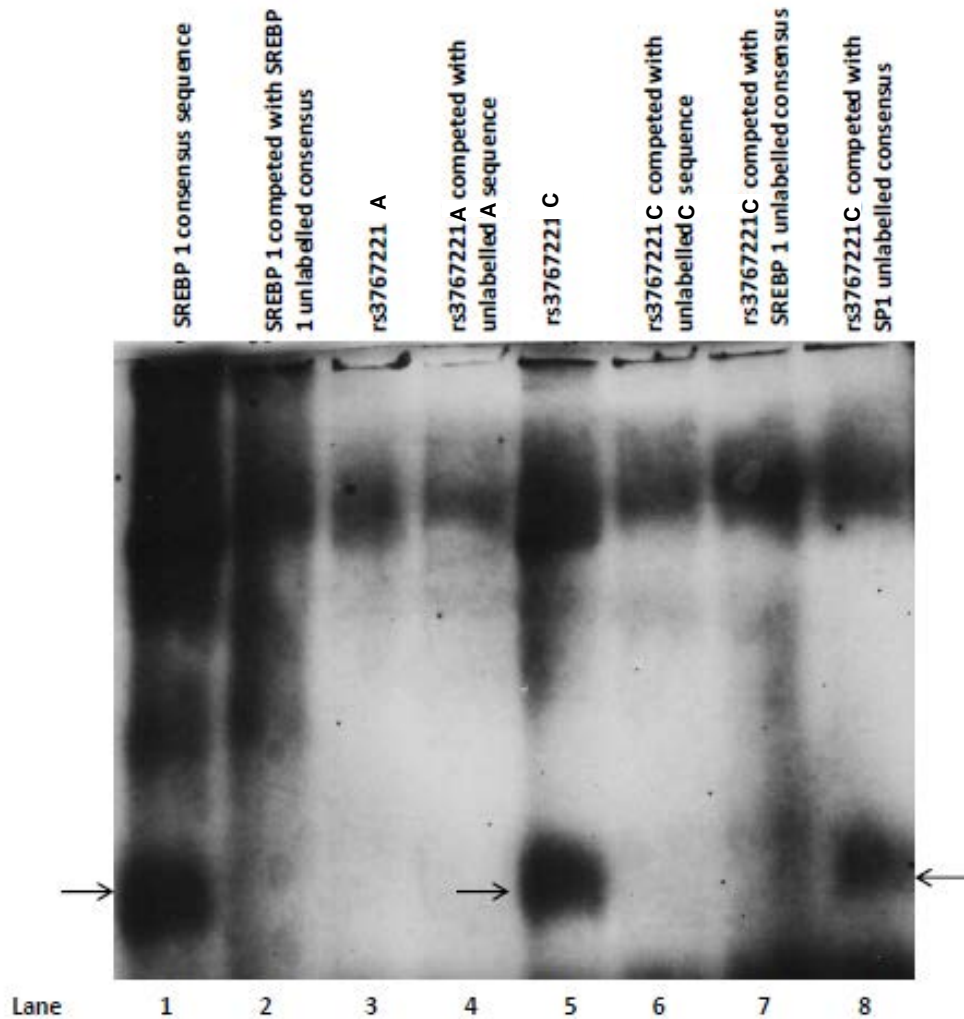


Figure 3.11. EMSA image for rs3767221; Lane 1: SREBP 1 specific binding as a positive control. Lane 2 shows the SREBP1 specific bands competed out by unlabelled SREBP1 consensus sequence. Lane 3 and 4 represent rs3767221 A (wild) with (Lane 4) and without (Lane 3) unlabelled A competitor with non-specific binding. Lane 5 shows rs3767221 C (rare) specific binding, which is competed out with an unlabelled rs3767221 C competitor in lane 6. This confirms differential binding between A and C alleles. Lane 7 shows rs3767221 C competed with unlabelled SREBP 1 and the rs3767221 C specific band is much less intense, suggesting SREBP 1 or a related transcription factor may bind at rs3767221 C. Lane 8 shows the additional band at rs3767221 G is not competed out by non-specific competitor Sp1.

3.2.6. TaqMan Gene Expression Results

3.2.6.1. TaqMan Gene Expression Results for the Analysis of rs11573156

A total of 67 samples from the DRD cohort (2.5) were genotyped using the TaqMan Genotyping Assay System (2.4.6.1) and the results for this are shown below in Figure 3.12. The Hardy-Weinberg equation showed no significant difference for this set of results. Semi-quantitative RT-PCR was then carried out on lymphocyte cell cDNA samples from 12 individuals of known genotype, across *PLA2G2A* exons 1 and 2 showed a trend towards lower levels in the rs11573156 CG and GG individuals, when compared to the 5 rs11573156 C homozygotes. The mean expression in the 7 combined rs11573156 CG and GG individuals showed ~4.9% lower levels of exon 1 and 2 compared to the homozygous C samples ($P = 0.09$) (Figure 3.13). Unlike exon 2, exons 5-6 are shown by Ensembl and NCBI databases to be present in all *PLA2G2A* transcripts and therefore expression measures for exon 5-6 were used as comparison for exon 1-2. For exons 5-6, no genotype effect on expression levels was seen ($P = 0.54$; Figure 3.13). These exon-specific gene expression assay results suggest that *PLA2G2A* exon 1 or 2, but not exons 5 and 6, are differentially expressed in a genotype-specific manner according to rs11573156. Each gene expression assay is measured relative to an endogenous control, *GAPDH* (Table 2.3)

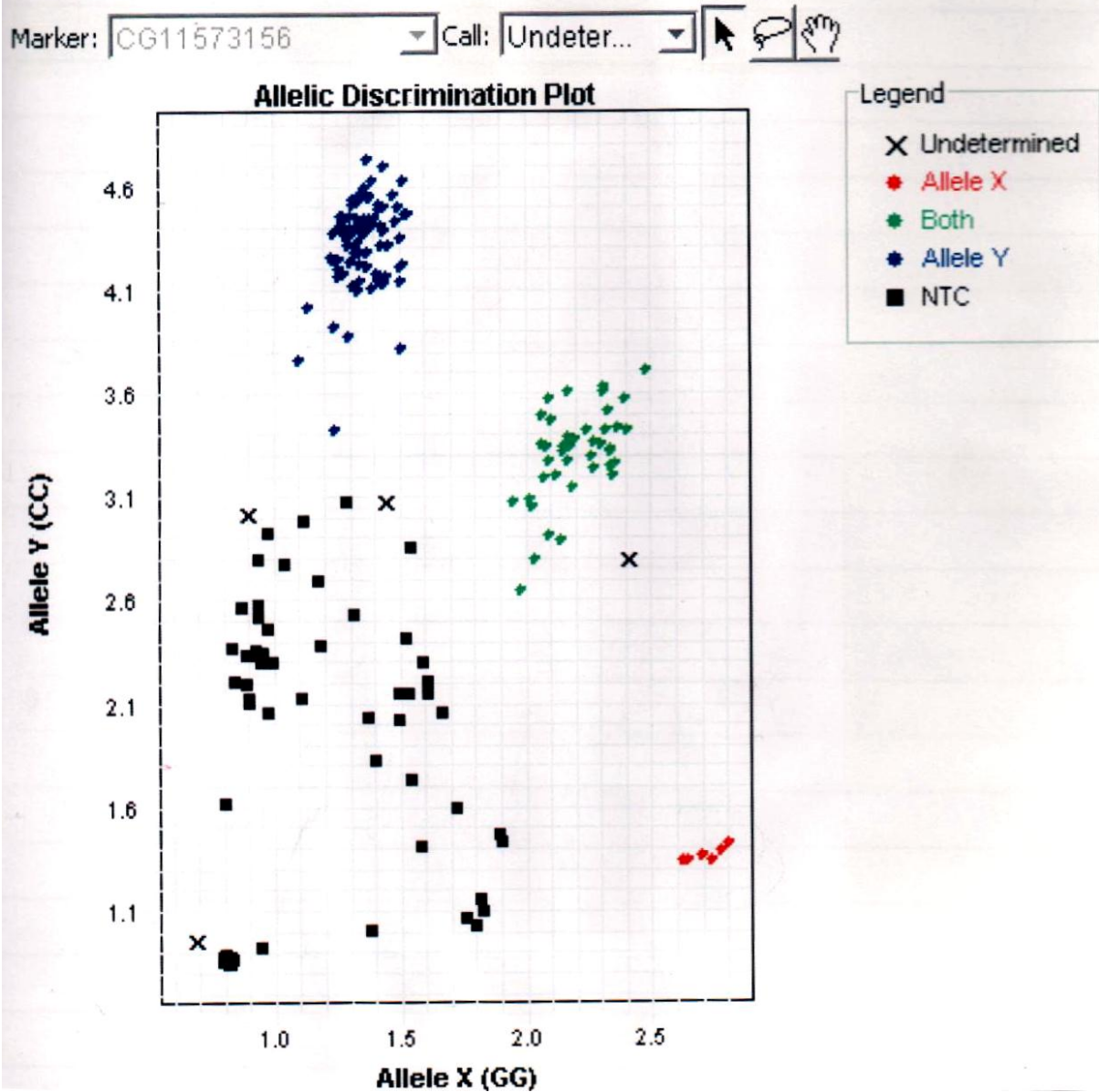
Allelic Discrimination Plot

Figure 3.12. The TaqMan Genotyping Assay results for rs11573156 in the DRD cohort of 67 samples. The autocaller function was used to differentiate between the genotypes. Black squares are NTCs. Black crosses are undetermined samples. Blue dots are designated CC common genotype. Green dots are designated CG heterozygous genotype. Red dots are designated GG the rare genotype. Despite there appearing to be some background noise (as suggested by the spread nature of the No Template Control (NTC) samples), the results only become invalid should any of the NTC samples record a measurement above the automatically determined thresholds to appear in any of the called groups, which none of these NTCs do.

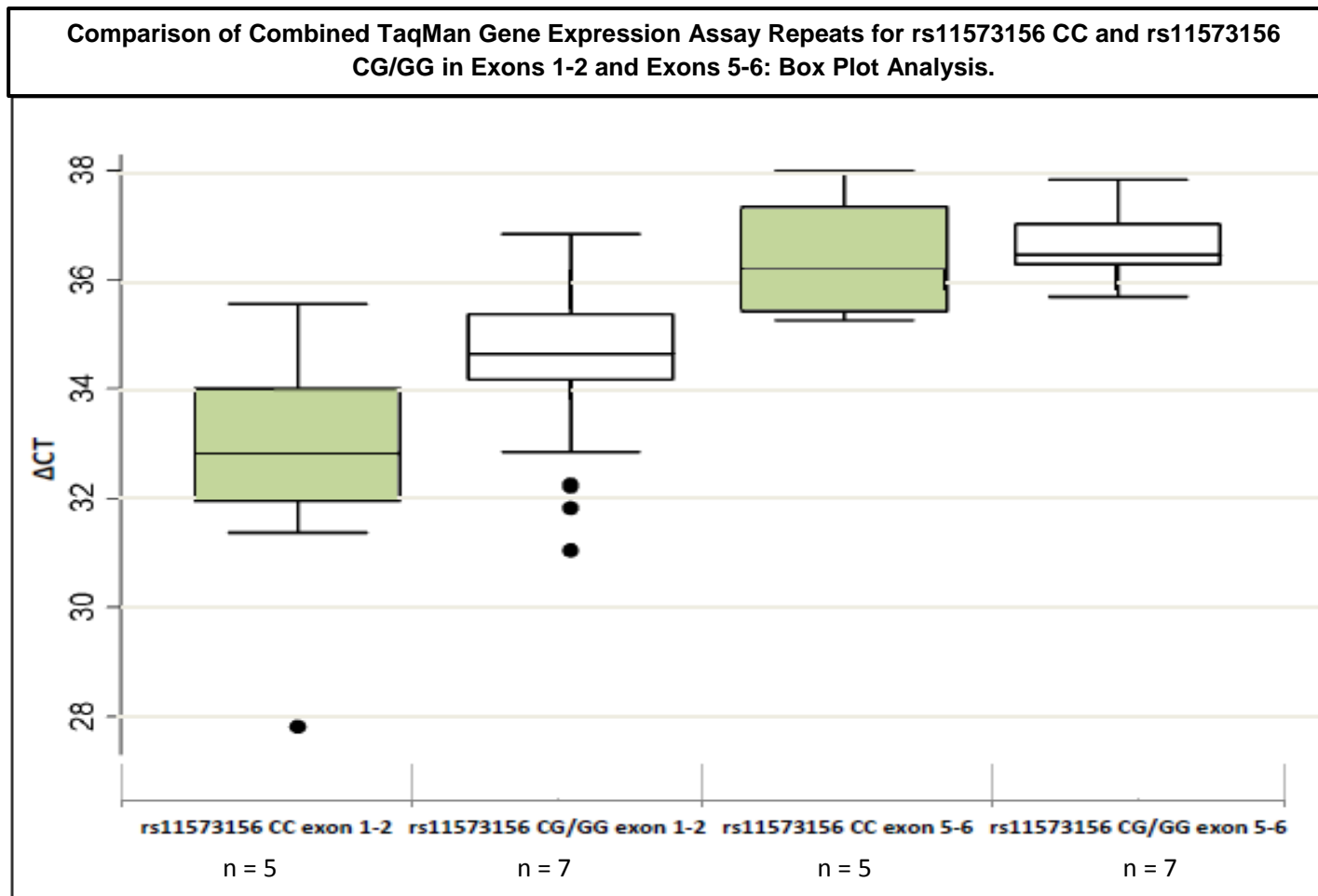


Figure 3.13. TaqMan gene expression assay results for rs11573156 C (Green) vs. rs11573156 G carriers (white) expression levels across *PLA2G2A* exon 1-2 and exon 5-6. All assays relative to endogenous control *GAPDH*. The X axis scale represents ΔC_t , the number of amplification cycles required to reach a pre-determined threshold of fluorescent signal. Error bars represent the SEM for each result. ● = Outlier. The black line through each box represents the median value of each dataset. It should be noted that the outliers for rs11573156 CG/GG exon 1-2 are not related to either the CC or the GG genotype specifically. A lower ΔC_t value corresponds to increased expression, since less cycles of amplification were required to achieve the manually determined threshold of expression (ΔR_n , see **2.4.9**. Figure 2.9-2.10b). The difference in expression between CC and CG/GG samples for exon 1-2 suggests that individuals with a CC genotype appear to have increased *PLA2G2A* exon 1-2 expression compared to individuals with a CC/GG genotype. There is no difference in expression between the same groups of individuals for *PLA2G2A* exons 5-6 suggesting that there is genotype specific differential expression in the region of *PLA2G2A* exons 1-2.

3.2.7. ASAP Study Gene Expression Analysis

Bioinformatics Study of rs10732279

The SNP rs11573156 was not present on the Illumina Human 610W-Quad Beadarray, so rs10732279 SNP (Figure 3.14) was used as a proxy ($R^2=0.91$ in Europeans, HapMap). It is therefore important to establish the likelihood of this SNP being functional rather than rs11573156.

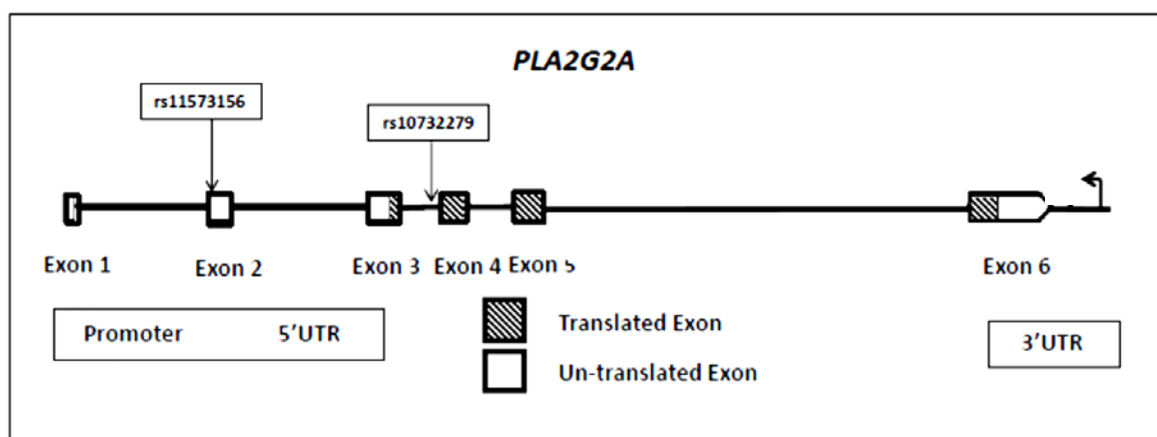


Figure 3.14 A schematic of the *PLA2G2A* gene. The proxy SNP rs11573156 is shown along with rs10732279. Rs10732279 is tagged by rs11573156 at $R^2=0.91$ and used as a proxy for this SNP in the ASAP study data since rs11573156 is not available on the Affymetrix Gene Chip. This figure also shows that rs10732279 is not present in the suspected alternate splice region where rs11573156 is shown.

The data from Ensembl and NCBI show that rs10732279 is positioned in an intronic region of *PLA2G2A* (NCBI Transcript ID: NM_000300.3). The allele frequencies for rs10732279 are A=0.76 and G=0.24 as shown on the NCBI SNP database (European population). According to the transcript and SNP data (UCSC and Ensembl), there are no known distinguishing features in the intronic region surrounding rs10732279 and it does not appear to be near an alternative splice site (Figure 3.3).

3.2.7.2. ASAP Study Gene Expression Analysis for rs11573156 Proxy

To examine potential allele-specific expression of *PLA2G2A*, we analysed *PLA2G2A* mRNA level data from the ASAP study (Folkersen, van't Hooft et al. 2010). Of the tissues investigated, liver, mammary arteries, dilated and non-dilated ascending aorta and heart, *PLA2G2A* mRNA was most highly expressed in the aortic adventitia, liver and heart (Figure 3.15) with the overall allele-specific differential expression of *PLA2G2A* mRNA was most pronounced in the liver (Figure 3.16). This suggests that a SNP or SNPs have the potential to regulate sPLA2-IIA expression in a CHD related tissue.

The proxy SNP rs10732279 was also the SNP that showed the greatest overall differential expression of *PLA2G2A* and explained 32% of the variance in *PLA2G2A* mRNA (Holmes et al. Submitted). The genotype effect of rs10732279 on *PLA2G2A* liver mRNA showed the strongest differential measurements of all the tissues examined and is presented in Figure 3.15; compared to the common A homozygotes, carriers of the rare G allele had roughly 25% higher expression of *PLA2G2A* ($P=1.67 \times 10^{-17}$). Expression of each exon compared by rs10732279 alleles was also examined and is shown below in Figure 3.17. Exon 2 was shown to be expressed at low levels compared to exons 3-6, suggesting that it could be alternately spliced, since it is not detected in as many samples as exon 3-6. Additionally, the difference in expression between rs10732279 AA and rs10732279 GG was significantly less in exon 2 than in exons 3-6. Again this could suggest that this is due to there being less G alleles being expressed in exon 2 since they are usually spliced out, whereas the A alleles maintain a regular expression level across all exons (Figure 3.17).

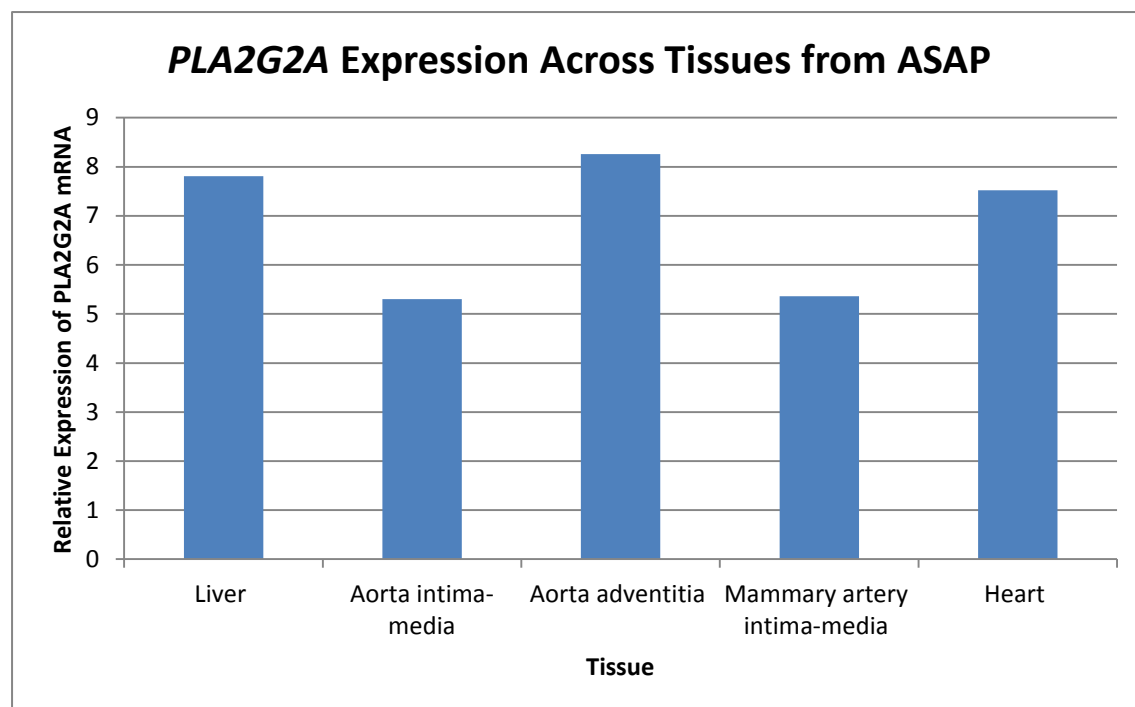


Figure 3.15 Expression of *PLA2G2A* mRNA measured across the tissues of the ASAP study; aorta media, aorta adventitia, liver, mammary artery and heart.

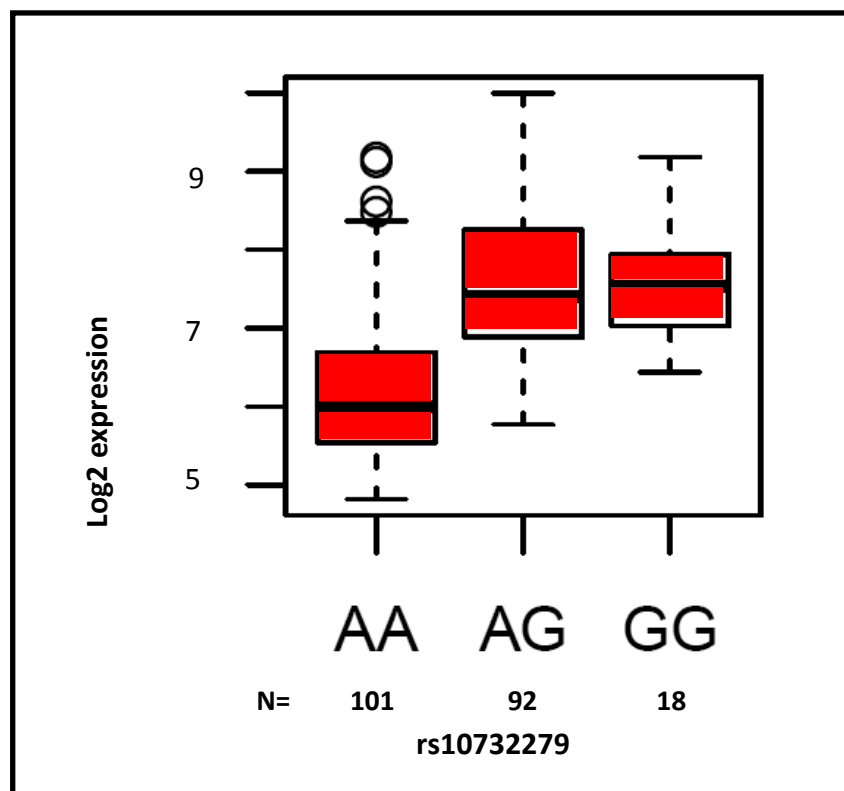


Figure 3.16. Results from the ASAP Study of the overall differential liver expression of *PLA2G2A* associated with rs10732279 (A>G) genotype. Results in the liver are showed since they are the strongest differential measurements regarding rs10732279. Rare G carriers have significantly higher total *PLA2G2A* expression than common A homozygotes ($P=1.67 \times 10^{-17}$).

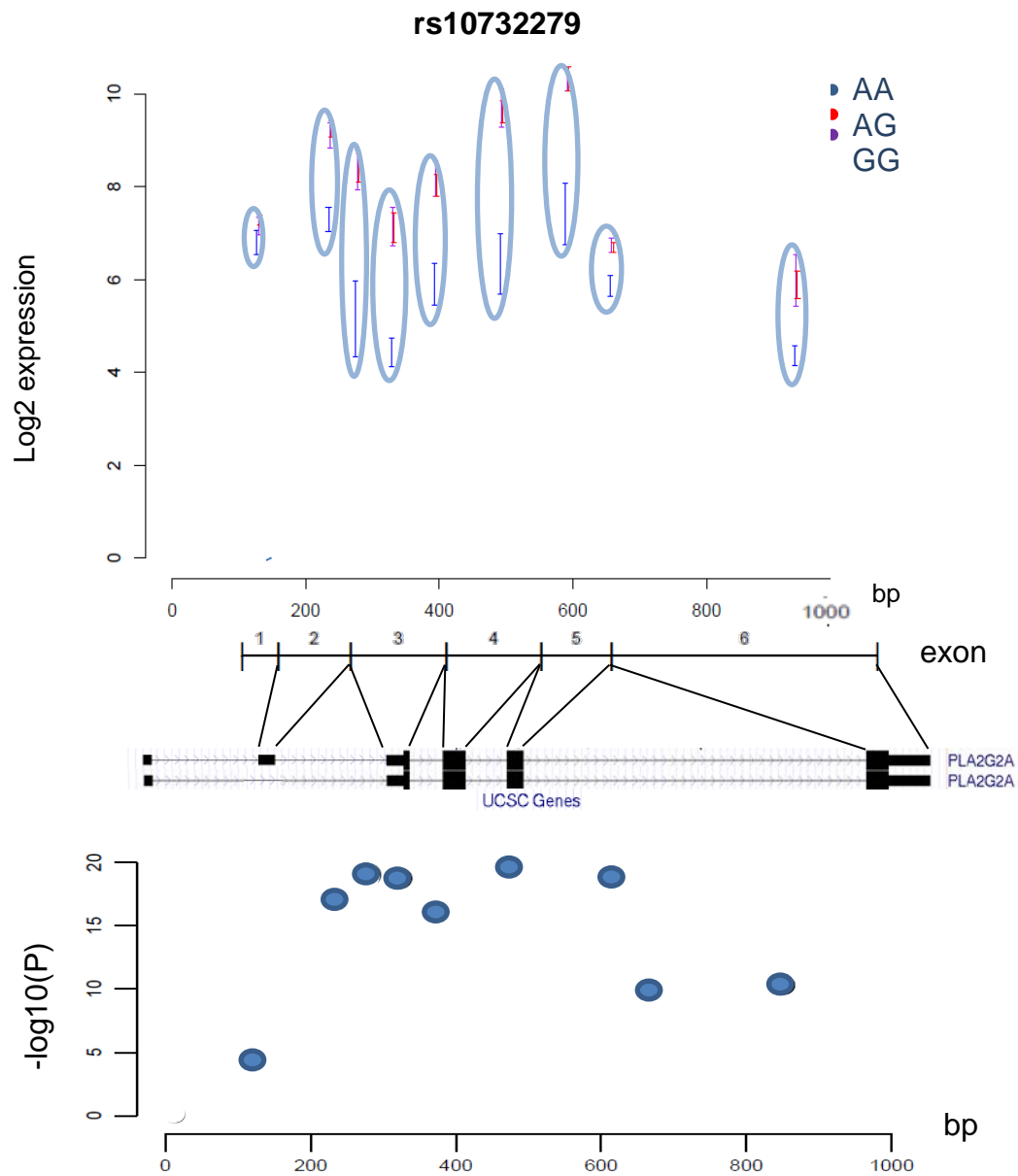


Figure 3.17. The top graph shows ASAP *PLA2G2A* Expression analysis for different rs10732279 genotypes. Results from the ASAP study; liver expression levels of exon-specific *PLA2G2A* probe sets stratified by rs10732279 genotype. *PLA2G2A* is covered by 9 exon-specific probe sets (1 in exon 2, and two each in exon 3-6)

The middle figure shows the known protein coding transcripts for *PLA2G2A*. Known transcribed *PLA2G2A* transcripts showing alternate exon 2 skipping from the UCSC Genome Browser (UCSC Genome Browser).

The bottom graph shows *PLA2G2A* Exon-Specific Expression in the Liver. Results from the ASAP study; Liver expression association between exon-specific *PLA2G2A* probe sets according to rs10732279 genotype. The y-axis give the $-\log_{10}(p\text{-value})$ for the *PLA2G2A* association for rs10732279, using an additive model. These values relate to the exon specific absolute expression levels shown in the top graph.

3.2.7.3. ASAP Study Gene Expression Analysis for rs3767221

Rs3767221 was not covered on the array nor was there a strong proxy for this SNP. However, using the 1000 genomes database Dr. Folkersen was able to impute rs3767221. The results predict an association of rs3767221 with *PLA2G2A*, albeit with less accuracy than for rs11573156. The mach1 R^2 value was shown to be 0.41. When the differential values of *PLA2G2A* for this imputed SNP were analysed they showed a significant association of $P=3.6 \times 10^{-4}$ (this was weaker than that seen for rs10732279, the proxy for rs11573156) with carriers of the common allele (A) showing higher expression compared to samples homozygous for the rare allele (C) (Figure 3.18).

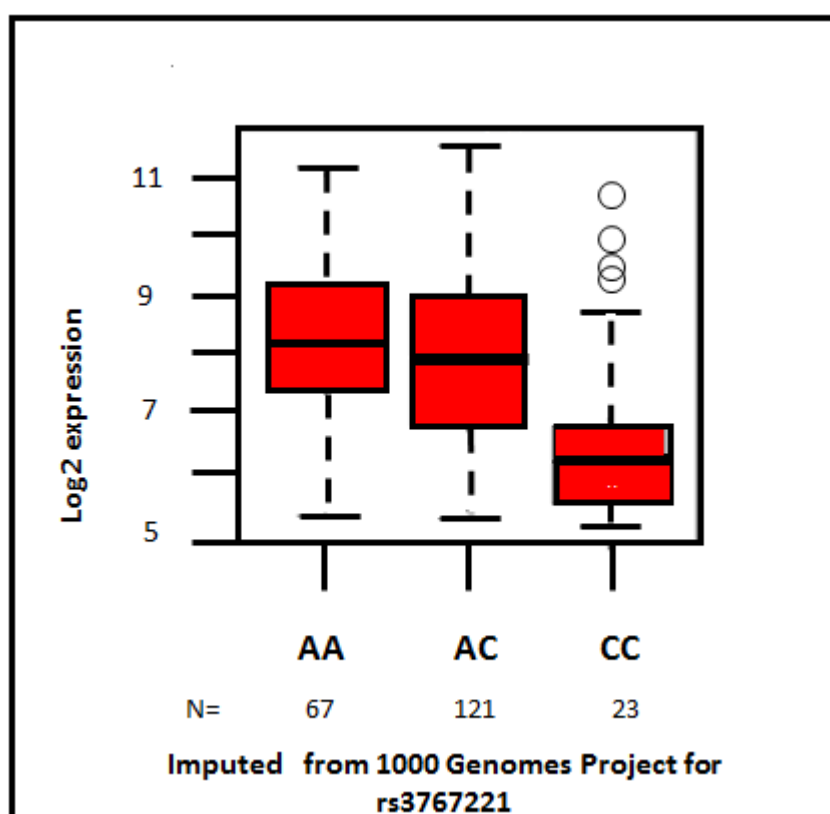


Figure 3.18. Results from the Advanced Study of Aortic Pathology relative to the 1000 Genomes Project imputed SNPs for *PLA2G2A*. The figure shows the predicted overall differential expression predicted for rs3767221 (A>C) in the liver. Common A carriers are predicted to have significantly higher *PLA2G2A* expression than rare C homozygotes, $P=3.6 \times 10^{-4}$.

3.3. Discussion

3.3.1. Rs11573156 C>G as a Functional SNP

It has been previously reported that the *PLA2G2A* 5'UTR and potential promoter SNP rs11573156 (in the region of the untranslated exon 2) and the 3'UTR SNP rs3767221, are independently associated with differences in levels of plasma sPLA2-IIA (Wootton, Drenos et al. 2006) and here I show that both influence gene expression by acting through different molecular mechanisms.

The difference in luciferase activity between the rs11573156 C and G constructs is modest in comparison to the baseline pGL3 basic measurement and of borderline statistical significance. Compared to pGL3 Basic alone (1RLU) the luciferase results were in the order of 1.2-1.6 RLU suggesting that this fragment of the promoter does not increase transcription. This is potentially explainable considering that the fragment that was analysed includes a previously identified negative regulatory region (Olivier, Fan et al. 1994; Fan, Paradon et al. 1997), and the inhibition of luciferase stimulation by a negative regulatory region in this fragment might explain the low luciferase activity measurements that we see. It should be noted that the rare (G) variant was introduced into the fragment using site directed mutagenesis and therefore both fragments were identical except for this one change in allele and therefore changes such as alternate splicing cannot be interpreted through these results. Bioinformatics analysis using the transcription factor binding prediction tool, MatInspector, indicated no potential, differential transcription factor binding at a matrix similarity of <0.95 in the region of rs11573156. A potential differential transcription factor binding site for Sp1 at a matrix similarity of <0.90 was predicted, but this was shown not to exist by EMSA analysis. Additionally, no predicted transcription factor binding was shown in the same region according to the UCSC genome browser. Predicted miRNA binding in the region of this SNP was also assessed using several known algorithms. Thus, in order to assess all potential transcription factor/miRNA binding in the region of rs11573156, I used the chemiluminescent EMSA technique. The lack of mobility shift in the EMSA corroborated my theory that there is no differential protein binding associated with this SNP since no specific differential protein binding between the rs11573156 C and G variants was seen, suggesting no differential protein binding in the region immediately surrounding rs11573156.

A clue that the genotype effect might be acting through another mechanism came from the observation that exon 2 is not always present in the reported transcripts of the gene (Ensembl reference sequence: ENST00000375111 and ENST00000400520). Considering the very low levels of luciferase activity in the luciferase experiment, it is likely that a negative regulatory element exists in the region of exon 2 in the 5'UTR/promoter region. Fan

et al, using deletion mapping of the *PLA2G2A* promoter in reporter gene assays, identified a negative element in the region -1969 to -1624, which incorporates exon 2 (Ensembl Genome Browser). When deleted from the full transcript, reporter gene expression increased by ~400% (Fan, Paradon et al. 1997).

To assess the relationship between rs11573156 and potential for differential alternate splicing, gene expression data was generated and analysed. Gene expression techniques analyse differences in mRNA or cDNA expression and therefore take into account post-transcriptional modifications such as splicing and allows us to see the effects of different SNPs and post-transcriptional modifications on gene expression and ultimately, protein expression. When *PLA2G2A* mRNA levels were tested in the limited samples of known genotype available to me, I observed that, compared to expression of exons 5 and 6, which showed no difference in expression by genotype, there was a non-significant lower levels of exons 1 and 2 with G carriers compared to C homozygotes. These results were confirmed by the results from the ASAP study. Since rs11573156 was not available on the ASAP arrays, rs10732279, a proxy for rs11573156 was studied across all the measured tissues and in liver was shown to be most significantly associated with a change in *PLA2G2A* mRNA levels with carriers of the rare C allele being significantly associated with increased total mRNA expression compared to the common A allele homozygotes in exons 3-6 but, importantly, the difference in expression between rs10732279 variants was significantly lower for exon 2. This suggests that lower mRNA expression associated with the rare G allele occurs with exon 2 due to the splicing out of this exon with this variant, so there is less of exon 2 being produced. In turn, the splicing out of exon 2 removes a negative regulatory region, thus increasing mRNA expression from all the following exons (3-6), showing an overall increase in total *PLA2G2A* mRNA associated with the rare C allele (Figure 3.17).

The trend for lower total expression of exon 1 and 2 seen with the rare G allele carriers compared to the C allele homozygotes in the TaqMan gene expression experiments therefore supports this data and theory, again suggesting that exon 2 could potentially be spliced out differentially with this allele and samples where exon 2 is spliced out will express lower mRNA in the exon 2 region. This splicing out of exon 2 would remove the negative regulatory region in this exon and would result in an increase of overall mRNA expression from exons 3-6, ultimately contributing to the regulation of sPLA2-IIA mass and sPLA2 activity which we know to be differentially associated with this allele from previous observational data (Wootton, Drenos et al. 2006). The TaqMan gene expression analysis that I carried out had limited power, since the sample of healthy lab volunteers was very small. I was also cautious of the fact that the TaqMan results were all at very high ΔCT values, suggesting that the assays were not very efficient. Therefore these results were used

as only a guide to indicate whether there might be a need for further testing. However, when taken together with the ASAP study, these results implied that, in the presence of the rare G allele, exon 2 was expressed less well, suggesting that it was preferentially skipped in the presence of the rare allele.

Rs10732279, which showed the most significant association with *PLA2G2A* expression in the ASAP study, is intronic and is not in close proximity to any predicted functional sites (UCSC Genome Browser). However our SNP of interest, rs11573156, for which it acted as a proxy, is in very close proximity (37bp) to a known exon skipping site (beginning at chr1:20306884).

To examine the possibility that rs11573156 is only marking the true functional SNP in linkage disequilibrium, the LD between rs11573156 and all known exon-boundary SNPs was examined using the 1000 Genomes data and the SNAP Pairwise LD program. There was no other genotyped variant at those positions showing LD with rs11573156 above an R^2 threshold of 0.7 paired.

3.3.2. Rs3767221 A>C as a Functional SNP

There is strong evidence to suggest that rs3767221 A>C in the 3'UTR of *PLA2G2A* is also functional. The rare allele was shown to be associated with reduced sPLA2-IIA mass compared to the wild type in observational studies (Wootton, Drenos et al. 2006). In this instance we have shown that this is likely to be due to differential protein binding between the wild type and rare variant, with a binding site created by the presence of the rare G allele. Lower luciferase activity was associated with the rare C variant, seemingly replicating the data from the observational studies that showed the rare C allele homozygotes to be associated with significantly lower sPLA2-IIA mass compared with the wild type A allele homozygotes in a per allele pattern. Additionally, EMSA analysis showed a specific protein bound differentially to the rare C allele fragment compared to the wild type A allele. Coincidentally, the migration of a specific band seen with rs3767221 C appeared to correspond partially to one of those of the SREBP 1 positive control, used for all EMSA work. Competition of the C allele with SREBP 1 unlabelled consensus sequence resulted in lower intensity of the specific band, which I attempted to confirm by supershift assay. However, I was unable to confirm SREBP 1 binding using a supershift, due to an inability to optimise the assay with the required thesis submission time frame. As previously shown, prediction by MatInspector suggests a potential cAMP binding protein site in the region of rs3767221 C that is not predicted for rs3767221 A and cAMP is known to effect human sterol regulation, including SREBP1 binding (Lu and Shyy 2006) and to play a role in sPLA2-IIA transcription regulation (Touqui and Alaoui-El-Azher 2001). The UCSC Genome Browser

identifies a dense potential transcription factor binding region in the sequence around rs3767221, which included SREBP1 and sPLA2 interacting transcription factors such as p300 (Giandomenico, Simonsson et al. 2003) and LXR alpha (Trousson, Bernard et al. 2009). Additionally the miRBase algorithm predicted a binding site in the region of rs3767221 and this binding site contains part of a known SREBP consensus sequence. Therefore, it may be that the protein binding differentially in the EMSA is related to SREBP, but further experiments must be carried out to confirm this.

Considering both the observational and luciferase assay results we can predict that the protein binding in this position in the 3'UTR of *PLA2G2A* is an inhibitor, which would correlate with the comparatively low luciferase activity and sPLA2-IIA mass associated with this SNP. Further experiments are needed to clarify the specificity of the protein beyond doubt and to identify the protein binding to this site.

Several likely potential miRNA binding sites were predicted by the miRNA prediction algorithms miRBase and PITA and the differential protein complex seen in the EMSA could also relate to a type of post-transcriptional modification such as an miRNA or a member of the ribonucleoprotein family, which are known to modify genes post-transcriptionally, by pre-mRNA splicing (Chaudhury, Hussey et al. 2011) and have been associated to phospholipid mechanisms in the cell (Maraldi, Mazzotti et al. 1992). Again, further investigation into what the differential protein complex is needs to be carried out.

Observational studies show higher sPLA2-IIA levels associated with the A variant compared to the C variant and this is concordant with the higher luciferase activity seen with the A allele and the higher *PLA2G2A* expression associated with A allele carriers predicted by the ASAP study, so we predict that the protein binding to this region in C allele carriers, would act as a repressor.

3.4. Summary of Findings

3.4.1. Summary of the Potential Functionality of rs11573156

These results provide strong evidence to suggest that rs11573156 is a functional SNP affecting the expression of *PLA2G2A*. The potential mechanism of action of rs11573156 is differential splicing of exon 2, where the rare G allele is associated with the splicing out of the untranslated exon 2 and creates a truncated 5-exon transcript from the dominant 6-exon transcript. This seems to affect the rate of transcription of sPLA2-IIA, potentially through the loss of a negative regulatory mechanism shown to be in this region.

3.4.2. Summary of the Potential Functionality of rs3767221

The data accumulated here also suggests that rs3767221 A>C is potentially functional SNP. However, this SNP appears to act through a mechanism of differential regulatory protein binding.

3.5. Conclusion

The major aim of this part of the study was to determine the functionality of these two SNPs in tissues relevant to CHD, with a view to validate their use as robust genetic instruments to be used in Mendelian Randomization (MR) analysis. The MR approach uses a genetic variant to validate the causal relationship between a biomarker and outcome, in this case sPLA2-IIA and CHD, by examining the relationship of the SNP with both biomarker and outcome and testing the concordance of effects (Hingorani and Humphries 2005).

The functional analysis of rs11573156 and rs3767221 showed both SNPs to have functional relationships regarding sPLA2-IIA mass and/or sPLA2 activity. Since rs11573156 appears to have the most significant association with sPLA2-IIA mass (Wootton, Arora et al. 2007) and activity (EPIC-Norfolk) we took this SNP forward to use for a Mendelian randomisation (MR) analysis in order to assess whether the relationship seen between sPLA2-IIA mass and CHD (Boekholdt, Keller et al. 2005; Mallat, Benessiano et al. 2007) is likely to be causal.

CHAPTER 4

Using Genetics to Test the Causality of sPLA2-IIA for CHD: A Mendelian Randomisation Approach

4. Background

The Mendelian randomisation data presented in this chapter is a result of a large collaboration of 32 studies. Thanks and credit is given to all of the Principal Investigators of each study and to all who worked on the genotyping and statistical analysis for each study. A list of Principal Investigators is provided in the acknowledgements section (p6).

4.1. Introduction

4.1.1. Formation of the sPLA2-IIA Mendelian Randomisation Study

In **Chapters 1** and **3** data from previous studies were examined and showed encouraging evidence supporting the hypothesis that there is causative role for sPLA2-IIA in CHD. The studies included both mechanistic analyses in animals (Ishimoto, Yamada et al. 2003; Webb, Bostrom et al. 2003; Tietge, Pratico et al. 2005; Bostrom, Boyanovsky et al. 2007; Koenig, Vossen et al. 2009) and observational studies in humans (Pruzanski, Keystone et al. 1988; Boekholdt, Keller et al. 2005; Mallat, Steg et al. 2005; Mallat, Benessiano et al. 2007; Koenig, Vossen et al. 2009; Oorni and Kovanen 2009). However, despite these promising findings it is not possible to definitively model a human disease process in animal models, and observational studies are susceptible to confounding and reverse causation.

In order to fully answer the question of the causation of sPLA2-IIA in the development of CHD, it is therefore necessary to take a different approach and this could also result in a better evaluation of the significance of sPLA2-IIA as a therapeutic target in CHD prevention. The approach used for this chapter took advantage of the rs11573156 C>G variant in the sPLA2-IIA encoding gene (*PLA2G2A*), that I had previously identified as a functional SNP (**Chapter 3**) and before that had been shown to be significantly associated with sPLA2-IIA mass (Wootton, Drenos et al. 2006). Using this SNP as a natural model of differential sPLA2-IIA mass/ sPLA2 activity the Mendelian randomisation principle was applied (Davey Smith and Ebrahim 2003). This variation is considered to be specific and unconfounded for the evaluation of the validity of sPLA2-IIA as a therapeutic target for the prevention of CHD. The concept of Mendelian randomisation and some examples of where it

could have been relevant and where it has been used successfully are explained in Chapter **1.7.3**.

4.1.2. The Effects of sPLA2-IIA Levels/sPLA2 Activity on Cardiovascular Risk Traits

There is a chance that sPLA2-IIA mass and sPLA2 activity act to effect CHD risk through actions on a known confounder. In order to assess this and to establish what traits should be adjusted for in the analysis of rs11573156 with risk of CHD, analysis was carried out in the EPIC-Norfolk population (a population described previously in **2.5.1**) to show how increasing tertiles of sPLA2-IIA mass and sPLA2 activity effect known cardiovascular risk factors (**4.2.2** Table 4.5A-B).

The cross-sectional association between log sPLA2-IIA mass and log sPLA2 enzyme activity was evaluated with the following traits: age, body mass index (BMI), C-reactive protein (CRP), systolic blood pressure (SBP), diastolic blood pressure (DBP), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), triglycerides (TG), apolipoprotein A1, apolipoprotein B, gender, type 2 diabetes (T2D) and smoking status. The tertiles of sPLA₂-IIA mass and sPLA₂ enzyme activity were created by dividing the population into 3 equal parts by either sPLA2-IIA mass or sPLA2 activity measures, with tertile 1 representative of the lowest measurements and tertile 3 the highest measurements. The mean and standard deviation (SD) for continuous traits and proportion of cases for binary traits for each tertile were calculated and tabulated. For sPLA2 activity; tertile 1 = 1.7-4.0 nmol/min/ml, tertile 2 = 4.0-4.9 nmol/min/ml and tertile 3 = 4.9-23.1 nmol/min/ml (Table 4.5A). For sPLA2-IIA mass; tertile 1 = 0.4-6.8 ng/ml, tertile 2 = 6.8-11.3 ng/ml and tertile 3 = 11.3-114.5 ng/ml (Table 4.5B). Next, a univariate linear and logistic regression analyses was completed for continuous and binary traits, respectively, using log(n) transformed sPLA2 IIA mass and sPLA2 enzyme activity as the explanatory (independent) variables. All analysis was carried out using Stata v11 (**2.4.11**).

4.1.3. Varespladib Literature Study

The known sPLA2 inhibitor, varespladib, was undergoing Phase III trials simultaneously with this Mendelian randomisation analysis was being undertaken and it was therefore important to compare the effects of the drug with the effect of the genetic instrument, rs11573156 C>G in order to contextualise the MR findings. According to the PRISMA guidelines (**2.4.1**) and using the search parameters outlined in **2.4.1** I was able to run a search that resulted in 4 primary human dose-response studies for varespladib (Figure 4.1). The four human based randomised trials identified (Table 4.1), totalled 1,300 participants (Rosenson, Hislop et al. 2009; Dzavik, Lavi et al. 2010; Rosenson, Hislop et al.

2010; Rosenson, Elliott et al. 2011). To compare the effect of the *PLA2G2A* SNP on sPLA₂-IIA mass and sPLA₂ enzyme activity in relation to pharmacological modification, I identified four randomized trials of the sPLA₂ inhibitor varespladib in a total of 1,300 individuals. A meta-regression of dose-finding studies indicated varespladib treatment produced a dose-dependent reduction in sPLA₂-IIA mass (P for meta-regression=0.06). The most frequently studied dose of varespladib (500mg/day) reduced sPLA₂-IIA mass by an average of 78% (95%CI: 62%, 94%). This effect is equivalent to approximately twice the effect seen for each C allele of rs11573156. The effect of varespladib on sPLA₂ enzyme activity was not reported in clinical trials since baseline activity was reported to be beneath the lower limit of quantification of the assay. No trial reported the effect of varespladib on sPLA₂-III, V or X mass.

Varespladib showed a strong association with sPLA₂-IIA mass. However, data extracted from these studies suggested no significant effects could be seen in the reduction of primary CHD risk and minimal effects were on secondary risk of heart disease (Rosenson and Gelb 2009; Dzavik, Lavi et al. 2010; Rosenson, Hislop et al. 2010; Rosenson, Elliott et al. 2011). This has been underpinned by the recent, premature termination of the phase III RCT, VISTA-16 (Nicholls, Cavender et al. 2011) due to reasons of futility. The outcome of this large scale MR analysis is therefore timely for helping to put the failure of this inhibitor trial into context.

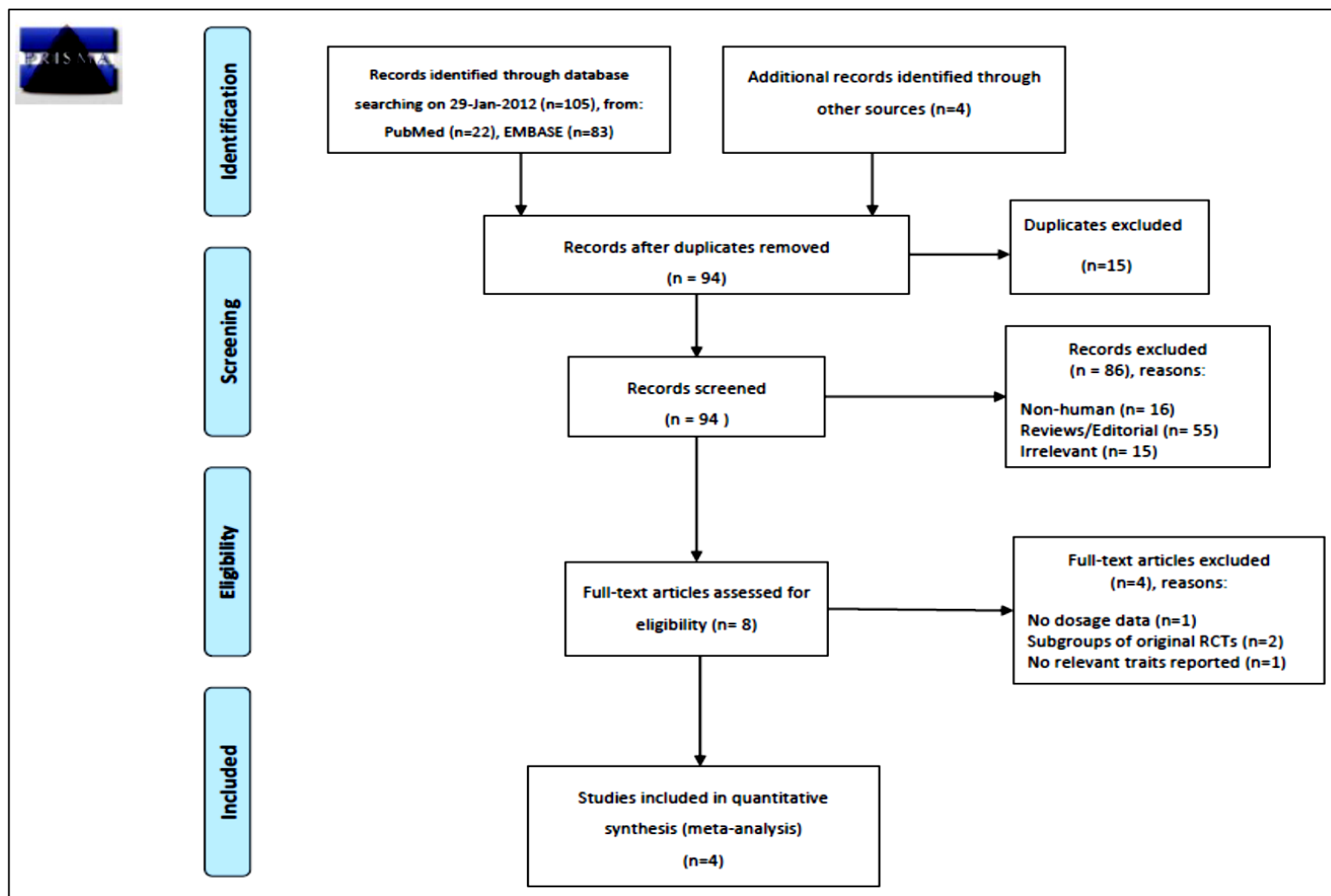


Figure 4.1. This flow diagram provided by PRISMA shows the process of the PRISMA guided analysis of varespladib randomised control trials. The pre-determined parameters resulted in 105 studies originally, which was narrowed down to 4 specific included studies by following the PRISMA guidelines for inclusion, demonstrated in the form of a flow diagram (<http://www.prisma-statement.org/statement.htm>).

Trial name	RCT phase	Years conducted	Clinical setting	Primary end-point	Age, mean (SD)	Sex, % female	Number randomised	Varespladib dose, mg/day (dosing frequency per day)	Trial duration, months
FRANCIS-ACS (Rosenson, Hislop et al. 2010)	IIB	2008-2010	ACS	Change in LDL-C from baseline to 8 weeks	59.05 (10.5)	25.3	625	500 (once)	6
PLASMA (Rosenson, Hislop et al. 2009)	II	2007	Stable CHD	Change in sPLA ₂ concentration or activity from baseline to 8 weeks	62 (11)	24	393	50/100/250/500 (twice)	2
PLASMA 2 (Rosenson, Elliott et al. 2011)	II	2007	Stable CHD	Change in sPLA ₂ concentration from baseline to 8 weeks	64 (12)	11	138	250/500 (once)	2
SPIDER-PCI (Dzavik, Lavi et al. 2010)	II	2007-2009	Stable CHD patients undergoing elective PCI	Elevation of cardiac enzymes (CK-MB, Troponin I) up to 24h PCI	63.4 (9.9)	12.5	144	500 (twice)	1

Table 4.1. Results from PRISMA guided analysis of varespladib randomised control trial studies. The table shows the design of each of the 4 trials identified from the literature search using the PRISMA guidelines (February 2012).

4.1.4. Identifying the Strongest Genetic Instrument

As presented in **Chapter 3**, two functional SNPs were identified in *PLA2G2A*; rs11573156 C>G and rs3767221 A>C. The SNP rs11573156 was chosen as the genetic instrument after investigation by genotyping in 5 studies with sPLA2-IIA mass and sPLA2 activity measures; including EPIC-Norfolk, GRACE-France, GRACE-Scotland, UDACS and FAST-MI (Populations defined in **2.5.1** Table 2.6) found evidence in favour of rs11573156 C>G being the stronger genetic instrument for the purposes of Mendelian randomisation analysis (results shown below in **4.2.1**). The statistical methods used to determine this are described in **2.4.11**. For each of the 5 studies, per C allele associations of each SNP with sPLA2-IIA mass and sPLA2 activity were calculated.

4.2. Results

4.2.1. Rs11573156 C>G the Strongest and Most Specific Genetic Instrument

The initial identification of rs11573156 C>G, along with 5 other tSNPs, showed significant association with sPLA2-IIA mass (Wootton, Drenos et al. 2006) in a Type II diabetes cohort, the University College London Diabetes and Cardiovascular Disease Study (UDACS) (Defined previously in Table 2.6 and Table 2.7). Further analysis of these 6 SNPs led to rs11573156 and rs3767221 being identified as the most highly associated with sPLA2-IIA mass and subsequently these two SNPs were genotyped in two additional studies. The European Prospective Investigation of Cancer (EPIC-Norfolk) (Day, Oakes et al. 1999) is a nested case-control prospective study based on 3,039 individuals sampled from general practices in Norfolk, England and the Global Registry of Acute Coronary Events (GRACE)-France is a study of 275 acute coronary syndrome (ACS) patients recruited from hospitals in France (Both of these studies are summarized in Table 2.6 and Table 2.7) (Investigators 2001). Both studies have measures for sPLA2-IIA mass and sPLA2 activity. The Hardy-Weinberg Equilibrium test for each SNP in each study is shown in Figure 4.2A-D. The baseline measurements for EPIC-Norfolk and GRACE France are shown in Table 4.3, categorised by prospective CHD cases and controls for EPIC-Norfolk and as acute coronary syndrome cases for GRACE. Additionally the association of genotype with all measured intermediate covariates is shown for both rs11573156 and rs3767221 in Table 4.4A-B.

SNP		EPIC-Norfolk				GRACE-France			
		N	sPLA2 activity	sPLA2-IIA mass	MAF	N	sPLA2 activity	sPLA2-IIA mass	MAF
rs3767221	11	1214	4.45 [3.85-5.22]	10.10 [6.80-16.05]	0.39	101	2.00 [1.50-2.70]	3.62 [2.15-8.67]	0.40
	12	1450	4.40 [3.85-5.15]	8.60 [5.90-12.90]		127	2.00 [1.40-2.90]	3.42 [1.66-5.71]	
	22	523	4.32 [3.81-4.97]	6.90 [5.10-9.10]		48	1.80 [1.20-3.30]	1.84 [1.13-3.18]	
	P value (trend)		0.06	7.91x10 ⁻⁴¹			0.50	0.0003062	
	P value ²		0.21	1.43x10 ⁻⁴⁷			0.36	0.0017	
rs11573156	11	1834	4.33 [3.78-5.04]	6.80 [5.08-9.0]	0.22	178	1.85 [1.20-2.80]	2.26 [1.37-4.24]	0.19
	12	1046	4.47 [3.90-5.23]	12.00 [8.74-16.47]		88	2.00 [1.50-3.20]	4.35 [2.61-1.25]	
	22	159	4.56 [3.91-5.39]	16.10 [10.94-23.73]		9	2.80 [1.50-2.80]	7.30 [4.66-18.84]	
	P value (trend)		0.00004132	1.08x10 ⁻¹⁷⁰			0.03	2.27x10 ⁻⁸	
	P value ²		0.0002	7.82x10 ⁻¹⁹³			0.07	2.48x10 ⁻⁶	

All results are represented as median (IQR). 11: homozygotes for the common allele; 12: heterozygotes; 22: homozygotes for the rare allele.

¹ unadjusted p values.

² age, gender and CHD adjusted p value.

Table 4.2. Shows the results of genotyping the *PLA2G2A* functional SNPs rs3767221A>C and rs11573156 C>G in the Epic Norfolk and GRACE France cohorts and their related associations with sPLA2 activity and sPLA2-IIA mass. The results show that rs11573156 has a stronger association to sPLA2 activity and sPLA2-IIA mass in both cohorts.

Stats for a 2-allele polymorphism				
SNP = rs11573156 C>G				
Observed				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	1834	1046	159	3039
	P	Q	p+q	
	0.776	0.224	1.000	
Expected				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	1828.1	1057.9	153.1	3039.0
Chi^2 test and level of significance				
	O	E	(o-e)^2/e	Chi^2 Sig level
	1834	1828.1	0.0	
	1046	1057.9	0.1	
	159	153.1	0.2	
			0.384	0.5353
Not sig at 1 df				
Notes:				
NTC = 72				
Undetermined = 287				

Stats for a 2-allele polymorphism				
SNP = rs11573156 C>G				
Observed				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	178	88	9	275
	P	Q	p+q	
	0.807	0.193	1.000	
Expected				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	179.2	85.6	10.2	275.0
Chi^2 test and level of significance				
	O	E	(o-e)^2/e	Chi^2 Sig level
	178	179.2	0.0	
	88	85.6	0.1	
	9	10.2	0.1	
			0.222	0.6378
Not sig at 1 df				
Notes:				
NTC = 106				
Undetermined = 3				

NTC= No template control; Undetermined = uncalled samples; 11 = homozygous for common allele; 12 = heterozygous; 22= homozygous for rare allele.

Figure 4.2A. The Hardy-Weinberg Equilibrium equation for rs11573156 C>G in EPIC-Norfolk. There is no significant difference between observed and expected values.

Figure 4.2B. The Hardy-Weinberg Equilibrium equation for rs3767221 A>C in GRACE France. There is no significant difference between observed and expected values.

Stats for a 2-allele polymorphism				
SNP = rs3767221 A>C				
Observed				
	rs3767221	rs3767221	rs3767221	
Genotypes	11	12	22	Total
Frequency	1239	1480	499	3218
	P	q	p+q	
	0.615	0.385	1.000	
Expected				
	rs3767221	rs3767221	rs3767221	
Genotypes	11	12	22	Total
Frequency	1217.0	1523.9	477.0	3218.0
Chi^2 test and level of significance				
	O	E	(o-e)^2/e	Chi^2 Sig level
	1239	1217.0	0.4	
	1480	1523.9	1.3	
	499	477.0	1.0	
				2.672 0.1021
Not sig at 1 df				
Notes:				
NTC = 72				
Undetermined = 166				

NTC= No Template Control; Undetermined = sample not called for genotype; 11= homozygous for the common allele; 12= heterozygous; 22= homozygous for the rare allele

Figure 4.2C. The Hardy-Weinberg Equilibrium equation for rs3767221 A>C in EPIC-Norfolk. There is no significant difference between observed and expected values.

Stats for a 2-allele polymorphism				
SNP = rs3767221				
Observed				
	rs3767221	rs3767221	rs3767221	
Genotypes	11	12	22	Total
Frequency	101	127	48	276
	p	q	p+q	
	0.596	0.404	1.000	
Expected				
	rs3767221	rs3767221	rs3767221	
Genotypes	11	12	22	Total
Frequency	98.0	132.9	45.0	276.0
Chi^2 test and level of significance				
	O	E	(o-e)^2/e	Chi^2 Sig level
	101	98.0	0.1	
	127	132.9	0.3	
	48	45.0	0.2	
				0.546 0.4600
Not sig at 1 df				
Notes:				
NTC = 106				
Undetermined = 2				

Figure 4.2D. The Hardy-Weinberg Equilibrium equation for rs3767221 A>C in GRACE France. There is no significant difference between observed and expected values.

EPIC-Norfolk				GRACE-France (n=276)	
Variable	Controls n=2237	Cases n=1138	P value	Variable	Mean (sd)
Age (years)	64.80 (7.70)	65.00 (7.90)	0.54	Age (years)	58.70 (13.00)
BMI (kg/m ²)	26.10 (3.40)	27.00 (3.70)	2.96x10 ⁻¹³	BMI (kg/m ²)	26.00 (4.40)
% Males (n)	63.10 (1411)	63.70 (725)	0.72	SBP (mm Hg)	135.00 (26.00)
SBP (mm Hg)	137.90 (17.60)	142.70 (18.80)	9.93x10 ⁻¹³	DBP (mm Hg)	76.90 (16.10)
DBP (mm Hg)	82.90 (10.90)	85.00 (11.90)	2.78x10 ⁻⁷	Cholesterol (mmol/L)	5.30 (1.30)
Cholesterol (mmol/L)	6.17 (1.11)	6.40 (1.21)	1.34x10 ⁻⁷	CRP (mg/L)	11.40 (13.80)
TG (mmol/L)	1.66 (0.83)	1.94 (0.96)	3.46x10 ⁻¹⁷	sPLA2 mass (ng/mL)	3.68 (4.22)
CRP (mg/L)	1.57 (1.83)	2.35 (2.75)	8.66x10 ⁻²¹	sPLA2 activity (nmol/min/mL)	3.09 (1.17)
sPLA2 mass (ng/mL)	8.72 (5.19)	10.00 (6.18)	5.75x10 ⁻¹⁰		
sPLA2 activity (nmol/min/mL)	4.36 (1.05)	4.69 (1.22)	1.53x10 ⁻¹⁵		

All EPIC Norfolk results are geometric mean (approx sd) except CRP, which is median (IQR). Sd = standard deviation

Table 4.3. Baseline measurements for the EPIC Norfolk and GRACE France cohorts. The measurements for EPIC-Norfolk are categorised by cases of CHD versus controls with adjustment for age and sex. The measurements for GRACE-France represent data from ACS patients.

EPIC-Norfolk Intermediate Traits By Genotype:										
rs3767221	N	BMI (kg/m ²)	SBP (mm Hg)	DBP (mm Hg)	CHOL (mmol/L)	TG (mmol/L)	LDL (mmol/L)	HDL (mmol/L)	CRP (mg/L)	LpPLA2 (U/L)
11	1214	26.5 (3.5)	139.4 (18.0)	83.6 (11.1)	6.25 (1.14)	1.78 (0.91)	4.07 (1.00)	1.27 (0.36)	1.6 (0.8-3.5)	49.4 (15.9)
12	1450	26.3 (3.5)	139.7 (18.1)	83.9 (11.5)	6.26 (1.18)	1.73 (0.87)	4.10 (1.03)	1.29 (0.36)	1.8 (0.8-4.0)	49.5 (15.6)
22	523	26.5 (3.5)	139.0 (18.1)	83.2 (10.9)	6.19 (1.11)	1.73 (0.86)	4.00 (0.98)	1.29 (0.38)	1.8 (0.8-3.9)	48.6 (15.7)
P value		0.64	0.73	0.95	0.30	0.31	0.30	0.39	0.13	0.61
rs11573156	N	BMI (kg/m ²)	SBP (mm Hg)	DBP (mm Hg)	CHOL (mmol/L)	TG (mmol/L)	LDL (mmol/L)	HDL (mmol/L)	CRP (mg/L)	LpPLA2 (U/L)
11	1834	26.4 (3.5)	138.9 (18.0)	83.4 (11.1)	6.23 (1.15)	1.75 (0.88)	4.06 (1.01)	1.28 (0.36)	1.7 (0.8-3.8)	49.3 (15.9)
12	1046	26.4 (3.5)	140.1 (18.2)	84.0 (11.4)	6.24 (1.18)	1.73 (0.87)	4.08 (1.05)	1.28 (0.36)	1.8 (0.8-3.9)	49.2 (15.8)
22	159	26.4 (3.1)	141.5 (19.5)	85.3 (12.3)	6.19 (1.06)	1.76 (0.85)	4.02 (0.88)	1.25 (0.38)	1.4 (0.7-2.9)	50.6 (15.5)
P value		0.64	0.02	0.03	0.78	0.67	0.82	0.35	0.03	0.63

All results are geometric mean (approx sd) except CRP, which is median (IQR)
Age and gender adjusted p values.

Table 4.4A. The Intermediate trait measurements of the Epic-Norfolk case/control cohort categorised by SNP genotype for the *PLA2G2A* SNPs rs11573156 C>G and rs3767221 A>C

GRACE-France Intermediate Traits By Genotype:						
rs3767221	BMI (kg/m ²)	SBP (mm Hg)	DBP (mm Hg)	CHOL (mmol/L)	LDL (mmol/L)	CRP (mg/L)
11 (n=101)	26.9 (4.4) 46	134.7 (24.7) 89	78.4 (14.4) 89	5.16 (1.09) 42	3.17 (0.96) 41	10.4 (3-23.3) 101
12 (n=127)	25.6 (4.2) 55	138.4 (25.5) 107	77.5 (17.4) 107	5.25 (1.31) 53	3.11(1.27) 47	8.7 (3-17.8) 126
22 (n=48)	25.5 (2.8) 13	127.1 (24.1) 37	74.8 (13.0) 37	5.72 (1.00) 17	3.78 (0.94) 17	9.2 (3.1-21.2) 48
P value		0.32	0.29	0.20	0.18	0.95
rs11573156	BMI (kg/m ²)	SBP (mm Hg)	DBP (mm Hg)	CHOL (mmol/L)	LDL (mmol/L)	CRP (mg/L)
11 (n=178)	26.2 (4.5) 77	134.5 (24.2) 157	78.3 (14.6) 157	5.32 (1.18) 77	3.27 (1.16) 72	9.2 (3-18) 177
12 (n=88)	26.1 (3.7) 36	135.3 (27.9) 68	75.0 (18.0) 68	5.27 (1.22) 32	3.20 (1.07) 30	10.4 (3.2-22.5) 88
22 (n=9)	26.3 (7.5) 2	145.0 (17.0) 7	81.2 (11.2) 7	4.83 (1.53) 4	2.92 (1.36) 4	5.5 (3-20.9) 9
P value		0.70	0.34	0.58	0.64	0.72

All results are geometric mean (approx sd) except CRP, which is median (IQR). Followed by Age and gender adjusted p values.

Table 4.4B The Intermediate trait measurements of the GRACE-France acute coronary syndrome cohort categorised by SNP genotype for the *PLA2G2A* SNPs rs11573156 C>G and rs3767221 A>C.

Upon analysis of the genotyped EPIC-Norfolk and GRACE-France cohorts I determined rs11573156 C>G as the most likely candidate for a strong genetic instrument to be used in Mendelian randomisation to study the role of sPLA2-IIA in the causation of cardiovascular events. The results showed rs11573156 to have a significantly strong association to both sPLA2-IIA mass ($P=1.08 \times 10^{-170}$) and sPLA2 activity ($P=4 \times 10^{-5}$). The SNP was also confirmed as a functional SNP as described by the functional analysis in **Chapter 3**.

Parallel to the genotyping of EPIC-Norfolk and GRACE-France, another French ACS cohort, FAST-MI was being genotyped for, among other SNPs, rs11573156 C>G. This genotyping and analysis was carried out by Dr Tabbassome Simon (APHP, Hôpital Saint Antoine, Department of Pharmacology – URC-EST; UPMC University, Site Saint Antoine, Department of Pharmacology) in a cohort of 973 individuals recruited from a nationwide ACS registry in France. The data indicated that rs11573156 C>G had the strongest association to sPLA2 activity and sPLA2-IIA mass. GRACE-Scotland has measures for both sPLA2-IIA mass and sPLA2 activity in 1,488 individuals with ACS, recruited from hospitals in Scotland. Again, genotyping data from this study has shown rs11573156 to be significantly associated to sPLA2-IIA mass and sPLA2 activity. Figure 4.3. gives an overview of this initial observational data. This analysis was carried out using Stata v11 and calculates the change in sPLA2-IIA mass and sPLA2 activity dependent on rs11573156 genotype. Figure 4.3 represents pooled meta-analysis estimates of the association between *PLA2G2A* rs11573156 with (A) sPLA2-IIA mass and (B) sPLA2 activity, separated by study setting into general populations (EPIC-Norfolk, UDACS) and acute coronary syndrome (GRACE-Scotland, GRACE-France, FAST-MI). The numbers in brackets represent the number of studies, participants, and between-study heterogeneity (measured by I^2).

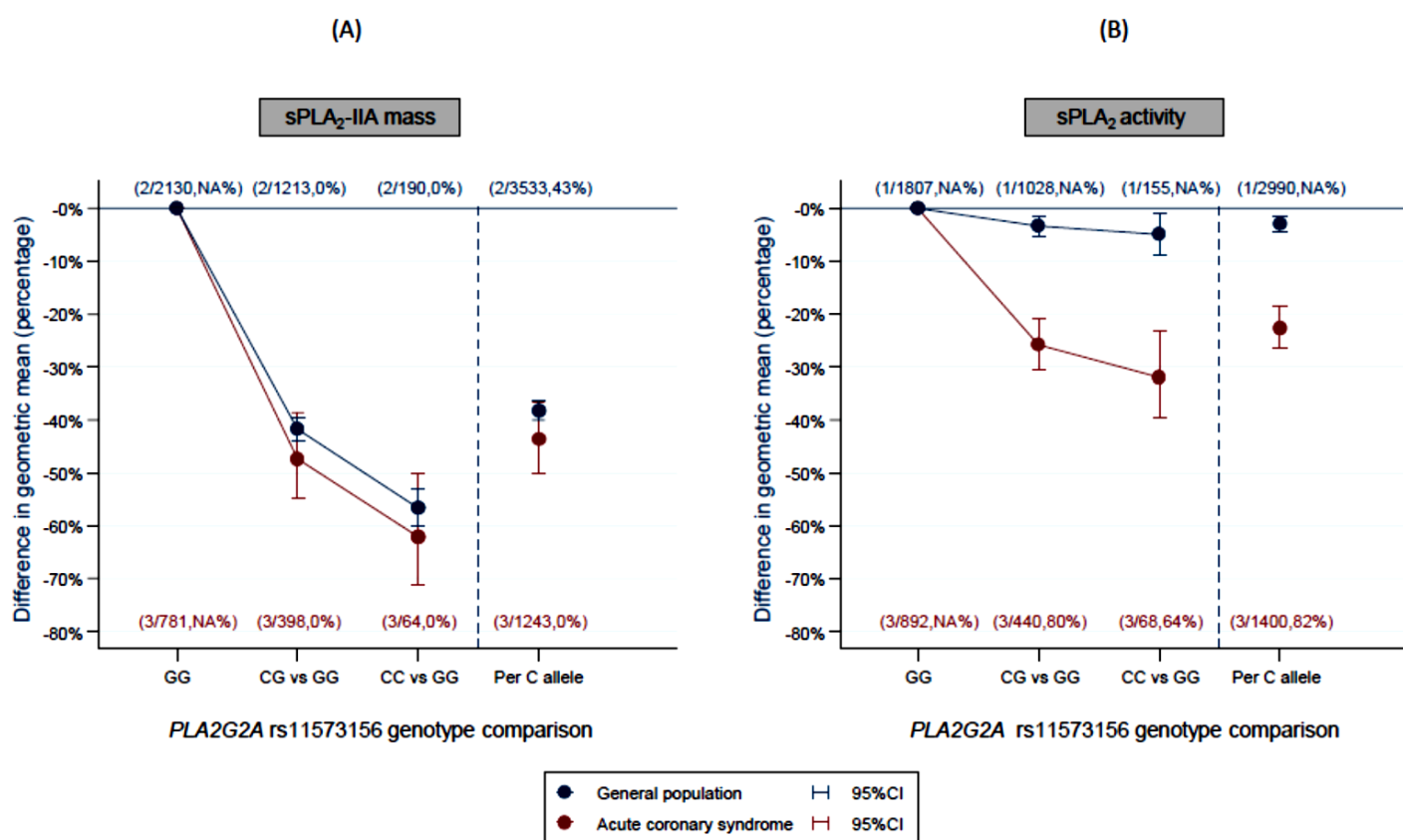


Figure 4.3. Schematic representations of rs11573156 C>G association with sPLA₂-IIA mass and sPLA₂ activity, based on 3 ACS studies; GRACE-France, GRACE-Scotland and FAST-MI (red) and 2 population based studies, UDACS and EPIC-Norfolk (blue). (A) Represents the association of the SNP with sPLA₂-IIA mass (B) Represents the association of the SNP with sPLA₂ Activity. Numbers in brackets represent the number of studies, participants, and between-study heterogeneity (measured by I^2) (Holmes et al. Submitted).

I requested some further analysis of the SNP in the ASAP cohort, which was carried out by Dr Lasse Folkersen as described in **Chapter 3**. The results suggested that *PLA2G2A* was most strongly expressed in the liver, aortic adventitia and heart tissues (Figure 3.15). *PLA2G2A* was additionally shown to have the highest total relative expression across the gene compared to *PLA2G5* or *PLA2G10* (log10 expression = 0.9 in *PLA2G2A* compared to 0.8 and 0.45 in *PLA2G5* and *PLA2G10* respectively). *PLA2G2A* mRNA was also seen to be differentially expressed at much higher levels than both *PLA2G5* and *PLA2G10* mRNA (Figure 4.4 A, B and C) in the liver. *PLA2G2A* SNPs showed no differential expression of *PLA2G5* or *PLA2G10*.

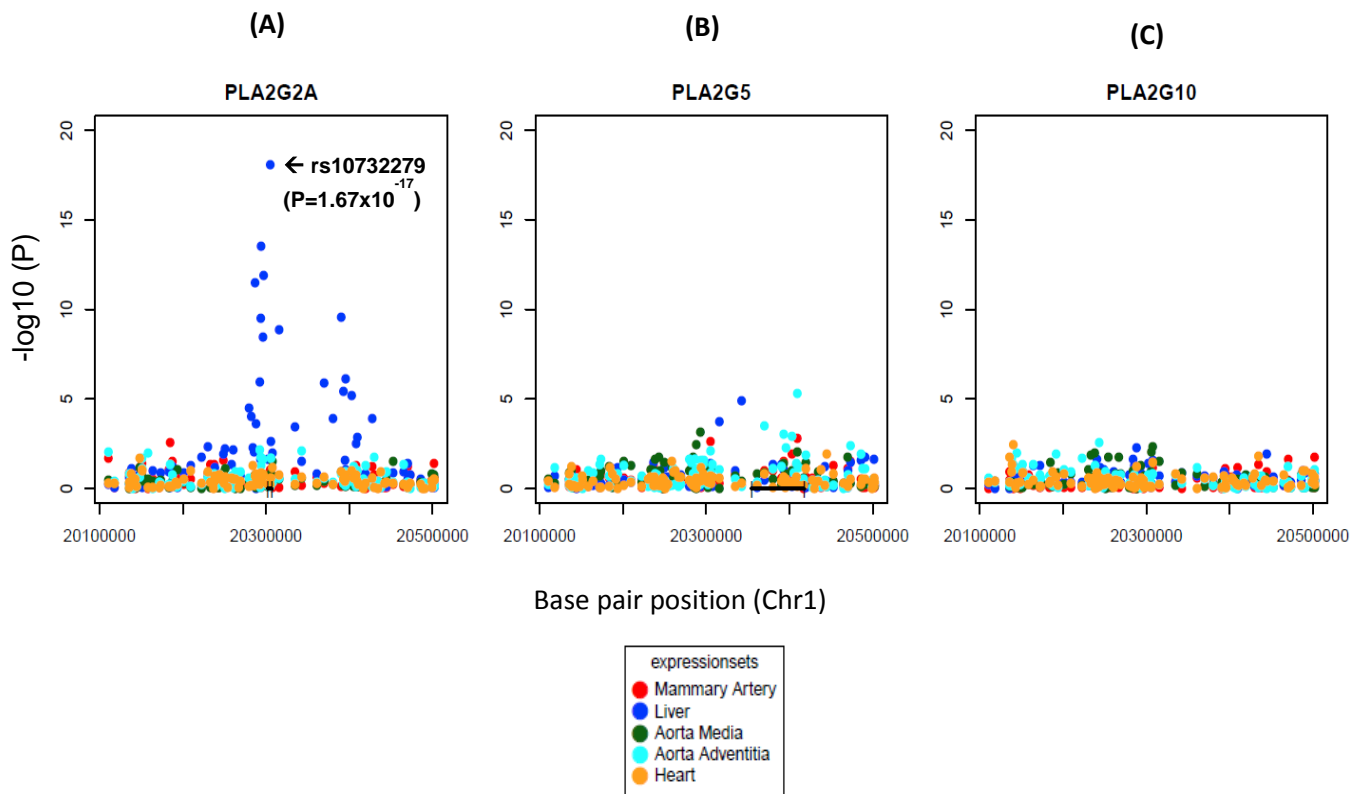


Figure 4.4A-C. Relative mRNA levels of *PLA2* Genes of interest associated with SNPs across these genes in a per allele model, measured in the tissue samples of the ASAP cohort. Results presented as Manhattan plots. (A) Relative mRNA levels in *PLA2G2A* (B) Relative mRNA levels in *PLA2G5* (C) Relative mRNA levels in *PLA2G10* (Holmes et al. Submitted).

As indicated in Figure 4.4A, the SNP that is clearly most strongly associated with *PLA2G2A* mRNA levels in the liver is rs10732279 ($P=8.71 \times 10^{-19}$). This is the rs11573156 proxy SNP that was identified in **Chapter 3** due to the strong LD between rs10732279 and rs11573156 ($R^2=0.91$ in European HapMap release 21). Since rs10732279 showed no association with either *PLA2G5* ($P=0.04$) or *PLA2G10* ($P=0.88$), the specificity of this SNP as a genetic instrument for the sPLA2-IIA isoform was confirmed.

As well as being a highly specific genetic instrument, data from the ASAP study showed (via its proxy, rs10732279) that rs11573156 explains 32% of *PLA2G2A* mRNA variability and these results have been replicated in an external data source of 206 transplant donor liver samples ($P=4.76 \times 10^{-8}$) (Innocenti, Cooper et al. 2011).

4.2.2. Results for Observational Analysis of sPLA2-IIA and Myocardial Infarction

Five studies (EPIC-Norfolk, UDACS, GRACE-France, GRACE-Scotland and FAST-MI), have sPLA2-IIA mass measurements, sPLA2 activity measurements, or both and rs11573156 was genotyped in each of these. Figure 4.2A and 4.2B show the genotyping data for EPIC-Norfolk and GRACE-France. The genotyping data for GRACE-Scotland and UDACS are shown below in Figure 4.5 and 4.6. The genotyping for FAST-MI was completed by one of the collaborating groups and I do not have access to the primary data. Initial analysis was for associations between sPLA2-IIA mass, sPLA2 activity and several cardiovascular risk factors in EPIC-Norfolk (Table 4.5A and Table 4.5B). These results showed that sPLA2-IIA mass and sPLA2 enzyme activity were associated with several established and emerging cardiovascular risk factors in the general population represented by EPIC-Norfolk. Generally an adjustment for cardiovascular risk factors reduced the association between sPLA2-IIA mass and sPLA2 enzyme activity with incident MI in the general population, though the association remained following multivariate adjustment.

Stats for a 2-allele polymorphism				
SNP = rs11573156 C>G				
Observed				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	1011	554	87	1652
	p	Q	p+q	
	0.780	0.220	1	
Expected				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	1004.2	567.6	80.2	1652
Chi^2 test and level of significance				
	O	E	(o-e)^2/e	Chi^2
	1011	1004.2	0.0	
	554	567.6	0.3	
	87	80.2	0.6	
				0.948
				0.3304
Not sig at 1 df				
Notes:				
NTC = 38				
Undetermined = 37				

NTC = No template control; Undetermined = Genotype call not assigned to sample; 11= homozygous for the common allele; 12 = heterozygous; 22= homozygous for the rare allele.

Figure 4.5. The Hardy-Weinberg Equilibrium equation for rs11573156 C>G in GRACE Scotland. There is no significant difference between observed and expected values at 1 degree of freedom.

Stats for a 2-allele polymorphism				
SNP = rs11573156				
Observed				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	357	183	24	564
	P	q	p+q	
	0.795	0.205	1.000	
Expected				
	rs11573156	rs11573156	rs11573156	
Genotypes	11	12	22	Total
Frequency	356.7	183.7	23.7	564.0
Chi ² test and level of significance				
	O	E	(o-e) ² /e	Chi ² Sig level
	357	356.7	0.0	
	183	183.7	0.0	
	24	23.7	0.0	
			0.008	0.9285
Not sig at 1 df				
Notes:				
NTC = 44				
Undetermined = 31				

NTC = No template control; Undetermined = Genotype call not assigned to sample; 11= homozygous for the common allele; 12 = heterozygous; 22= homozygous for the rare allele.

Figure 4.6. The Hardy-Weinberg Equilibrium equation for rs11573156 C>G in UDACS. There is no significant difference between observed and expected values at 1 degree of freedom.

Tertile (min-max, nmol/min/ml) of sPLA ₂ activity										
	1 (1.7-4.0)			2 (4.0-4.9)			3 (4.9-23.1)			
Continuous traits	N	Mean	SD	N	Mean	SD	N	Mean	SD	P value §
Age (yrs.)	1107	64.19	7.84	1107	64.54	7.89	1107	65.80	7.48	7.17x10-07
BMI (kg/m ²)	1107	26.02	3.44	1106	26.61	3.75	1104	27.32	3.71	1.00x10-15
CRP* (mg/dl)	1086	0.41	1.2	1087	0.53	1.16	1090	0.83	1.14	1.87x10-16
Systolic BP (mmHg)	1103	138.62	18.33	1105	139.76	18.13	1106	143.71	18.01	7.46x10-10
Diastolic BP (mmHg)	1103	83.22	11.34	1105	83.88	11.02	1106	86.03	11.69	2.57x10-07
HDL-C (mmol/l)	1088	1.38	0.4	1070	1.35	0.4	980	1.28	0.38	2.99x10-08
LDL-C (mmol/l)	1088	3.93	0.94	1070	4.14	0.99	981	4.37	1.1	4.26x10-21
TG* (mol/l)	1091	0.29	0.39	1093	0.52	0.44	1092	0.87	0.5	5.40x10-168
Apolipoprotein A1 (mg/dl)	976	159.74	28.98	956	160.62	29.45	955	159.18	29.93	0.94
Apolipoprotein B (mg/dl)	1038	120.49	26.91	1024	130.77	29.12	993	146.7	35.2	5.61x10-72
Binary traits	N	Proportion	SD	N	Proportion	SD	N	Proportion	SD	P value ¢
Female sex (%)	1107	33.06	0.47	1107	35.95	0.48	1107	41.73	0.49	1.27x10-05
T2D (%)	1107	2.62	0.16	1105	4.43	0.21	1106	3.35	0.18	0.10
Ever smoker (%)	1095	60.00	0.49	1098	59.56	0.49	1092	67.58	0.47	2.42x10-04

Footnotes Abbreviations: BMI: body mass index; CRP: C-reactive protein; HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; T2D: type 2 diabetes

* log-transformed (units do not apply); § derived from univariate linear regression using log sPLA₂ activity as the independent variable; ¢ derived from univariate logistic regression using log sPLA₂ activity as the independent variable.

Table 4.5A. The association between tertiles of sPLA₂ activity and traditional cardiovascular risk factors in 3,371 individuals in EPIC-Norfolk (Holmes et al. Submitted).

Tertile (min-max, ng/ml) of sPLA ₂ -IIA mass										
	1 (0.4-6.8)			2 (6.8-11.3)			3 (11.3-114.5)			
Continuous traits	N	Mean	SD	N	Mean	SD	N	Mean	SD	P value §
Age (yrs.)	1125	63.51	8.24	1124	64.93	7.59	1122	66.07	7.17	1.10x10 ⁻¹⁷
BMI (kg/m ²)	1125	26.20	3.28	1120	26.69	3.66	1122	27.03	3.99	5.46x10 ⁻⁰⁹
CRP* (mg/dl)	1100	0.22	1.05	1108	0.53	1.15	1104	1.01	1.21	1.67x10 ⁻⁷⁴
Systolic BP (mmHg)	1122	139.54	17.85	1122	140.63	17.64	1120	141.83	19.21	4.36x10 ⁻⁰³
Diastolic BP (mmHg)	1122	84.13	11.08	1122	84.49	11.20	1120	84.51	11.99	0.69
HDL-C (mmol/l)	1073	1.31	0.38	1051	1.35	0.40	1059	1.36	0.40	0.03
LDL-C (mmol/l)	1073	4.05	0.98	1052	4.17	1.03	1059	4.20	1.05	0.01
TG* (mol/l)	1114	0.57	0.50	1106	0.56	0.51	1105	0.55	0.51	0.06
Apolipoprotein A1 (mg/dl)	977	156.23	28.26	972	160.14	28.58	972	163.27	31.02	9.03x10 ⁻⁰⁶
Apolipoprotein B (mg/dl)	1029	129.79	32.28	1024	133.38	31.28	1038	134.31	33.42	0.07
Binary traits	N	Proportion	SD	N	Proportion	SD	N	Proportion	SD	P value †
Female sex (%)	1125	21.33	40.98	1124	35.68	47.93	1122	53.21	49.92	8.94x10 ⁻⁵⁹
T2D (%)	1124	3.20	17.62	1123	3.03	17.14	1120	4.11	19.85	0.06
Ever smoker (%)	1116	62.01	48.56	1106	62.30	48.49	1113	62.80	48.35	0.78

Footnotes Abbreviations: BMI: body mass index; CRP: C-reactive protein; HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; T2D: type 2 diabetes

* log-transformed (units do not apply); § derived from univariate linear regression using log sPLA₂-IIA as the independent variable; † derived from univariate logistic regression using log sPLA₂-IIA as the independent variable.

Table 4.5B. The association between tertiles of sPLA₂-IIA mass and traditional cardiovascular risk factors in 3,371 individuals in EPIC-Norfolk (Holmes et al. Submitted).

Analysis of the relationship between sPLA2-IIA mass (Figure 4.7A) and sPLA2 activity with incident/non-fatal myocardial infarction risk was examined by creating frequency distribution plots (4.6A-B) using Stata v11, which showed that both relationships followed a log-linear association and the likelihood ratio test P-value to support a quadratic over a linear model was 0.77 for sPLA2-IIA mass and 0.90 for sPLA2 activity. A similar association was identified between sPLA2-IIA mass with recurrent fatal/non-fatal myocardial infarction events in subjects with ACS in the combined GRACE-France and GRACE-Scotland cohorts (Figure 4.7B). After adjustment for multiple variables including age, sex, smoking, blood pressure and LDL-cholesterol, the odds ratio (OR) per one standard deviation (SD) reduction of sPLA2-IIA mass and sPLA2 activity with incident fatal/nonfatal myocardial infarction was 0.84 (95%CI: 0.77,0.91) and 0.86 (95%CI: 0.79, 0.94), respectively. The log odds of CHD and associated 95% confidence intervals were derived from a predicted linear model.

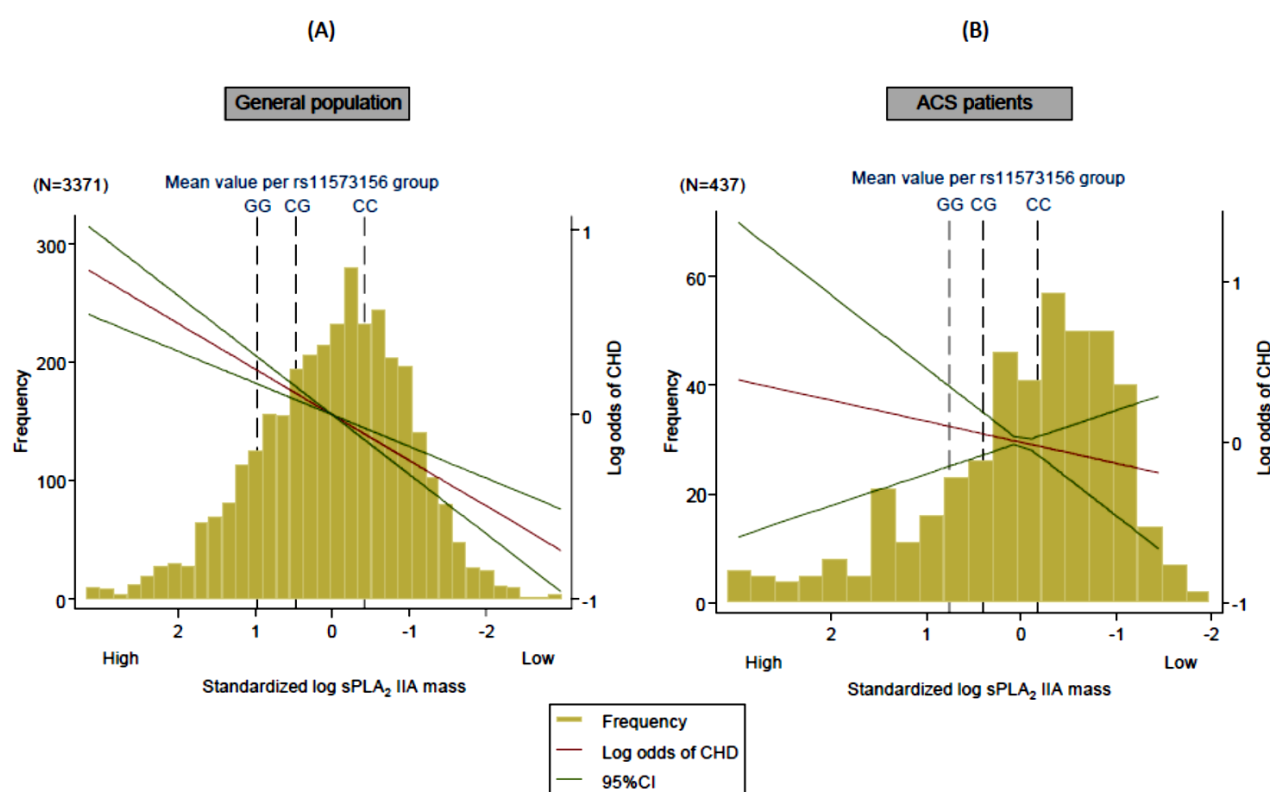


Figure 4.7. The frequency distribution of sPLA2-IIA mass in: (A) a general population study (EPIC-Norfolk); and (B): two ACS cohorts (GRACE-France and GRACE-Scotland), with a risk of (A) incident and (B) recurrent fatal/non-fatal myocardial infarction (red and green lines derived from predicted linear model adjusted for age and sex). Black dashed vertical lines represent the mean sPLA2-IIA mass per genotype (Holmes et al. Submitted).

4.2.3. Mendelian Randomisation Results: rs11573156 and Cardiovascular Outcomes

The studies used for Mendelian randomisation analysis are described in **2.5** in Table 2.6. and Table 2.7. Of these studies I genotyped IMPROVE, TPT, CYPRUS, NPHSII and EAS by TaqMan genotyping, a total of 11,534 samples. I also analysed this data using the Hardy Weinberg equilibrium and found none of the results were significant at 1 degree of freedom and that there was no significant difference in the allele frequencies across all of these studies.

Major vascular events (MVE) were categorised based on the outcomes described by each study, which were then standardised to fatal/nonfatal MI or stroke. For incident MVE (fatal/nonfatal MI or stroke) association with rs11573156, in 13 population studies comprising 8,021 incident MVE in 56,359 individuals, the summary per C allele OR was 1.02 (95%CI: 0.98, 1.06). The summary OR for individual outcomes were: incident nonfatal myocardial infarction, 1.04 (95%CI: 0.98, 1.10); incident nonfatal stroke, 1.00 (95%CI: 0.93, 1.07) and incident fatal myocardial infarction/stroke, 1.01 (95%CI: 0.93, 1.10) (Figure 4.8). These results were adjusted for age, sex, CRP, SBP, BMI, smoking status, T2D, TG and LDL-cholesterol. A power calculation justifying the population numbers used in these meta-analyses studies is described in **2.5.1**.

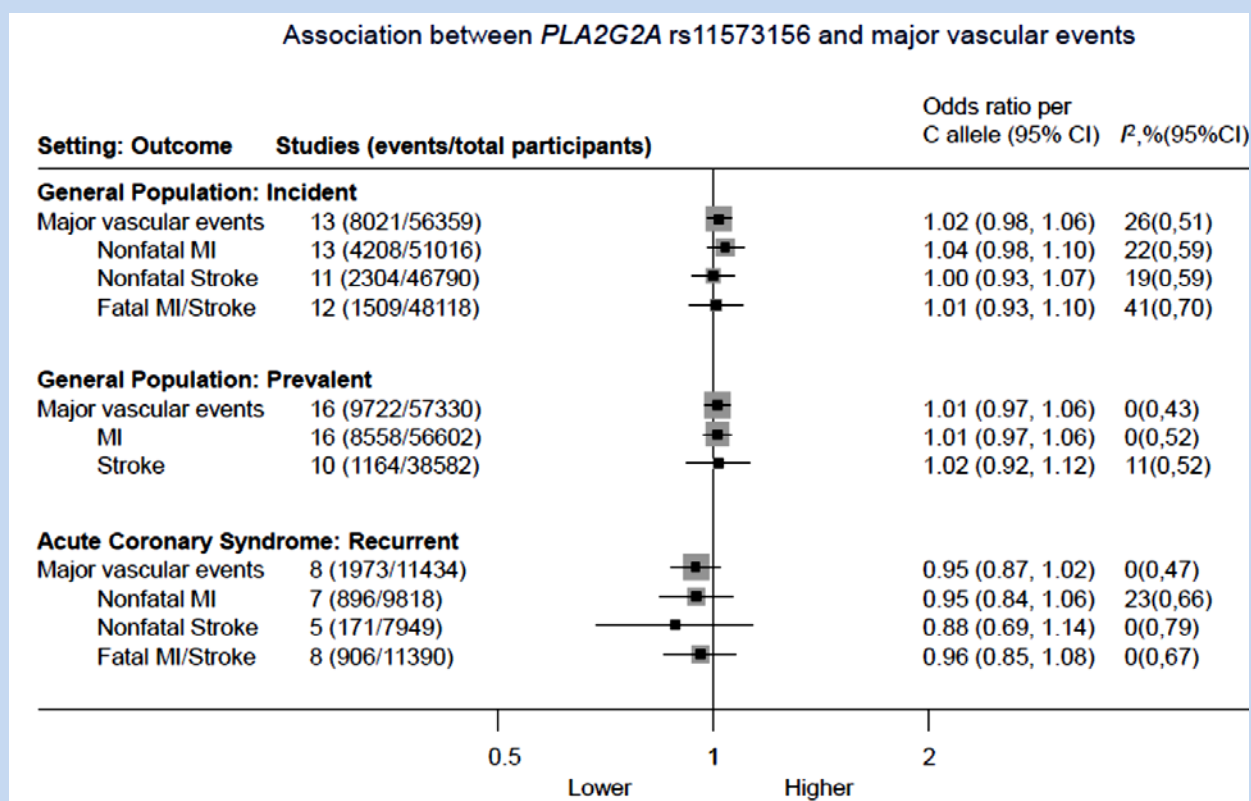


Figure 4.8. Meta-analysis pooled estimates of association between rs11573156 and major vascular events (including individual components) stratified by clinical setting into general populations and ACS patients. Adjusted for age, sex, CRP, SBP, BMI, smoking status, T2D, TG and LDL. Each plot represents the rs11573156 per C allele OR, with genotype grouping arranged to mimic the effects of pharmacological lowering of sPLA2-IIA so that if sPLA2-IIA mass were to reduce risk of cardiovascular events the OR should be less than one. In this instance MVE= fatal/nonfatal myocardial infarction or stroke. I^2 represents the percentage heterogeneity across the meta-analysis.

For prevalent MVE (MI or stroke) association with rs11573156, in 16 population studies comprising 9,722 prevalent MVE in 57,330 individuals, the summary per C allele OR was 1.01 (95%CI: 0.97, 1.06). The summary OR for individual outcomes were: prevalent myocardial infarction, 1.01 (95%CI: 0.97, 1.06); prevalent stroke, 1.02 (95%CI: 0.92, 1.12) (Figure 4.8).

The association between rs11573156 and recurrent cardiovascular events in patients with acute coronary syndrome was assessed in 8 studies (1,973 recurrent MVE in 11,434 participants with ACS). Results showed a summary per C allele OR for rs11573156 with recurrent MVE (fatal/nonfatal MI or stroke) of 0.95 (95%CI: 0.87, 1.02). The summary OR for individual components were: nonfatal MI 0.95 (95%CI: 0.84, 1.06); nonfatal stroke 0.88 (95%CI: 0.69, 1.14), and; fatal MI/stroke 0.96 (95%CI: 0.85, 1.08). These findings were consistent in a random-effects meta-analysis (Figure 4.8).

Additionally, in the study set in patients with established vascular disease or high cardiovascular risk (SMART), the per C allele OR of rs11573156 with recurrent MVE (fatal/nonfatal MI or stroke) was 0.92 (95%CI: 0.80, 1.07). For the study set in patients with ACS and undergoing PCI, the per C allele OR of rs11573156 with coronary artery restenosis was OR 1.04 (95%CI: 0.82, 1.31).

In order to draw a definitive conclusion from the Mendelian randomisation, the observed effects odds ratios (as seen above) had to be compared with the expected outcome. The expected association between rs11573156 and incident cardiovascular events was calculated assuming a causal relationship between circulating sPLA₂-IIA mass and sPLA₂ activity with fatal/nonfatal myocardial infarction. We estimated the expected per C allele OR of rs11573156 with fatal/nonfatal MI events. For sPLA₂-IIA mass, the expected genetic association was OR 0.88 (95%CI: 0.82, 0.94) in general populations, with a similar outcome predicted for ACS studies. Compared to the observed associations, using rs11573156, for either sPLA₂ activity or sPLA₂-IIA mass with fatal/nonfatal myocardial infarction, the expected associations showed very little overlap in all study settings (Figure 4.9).

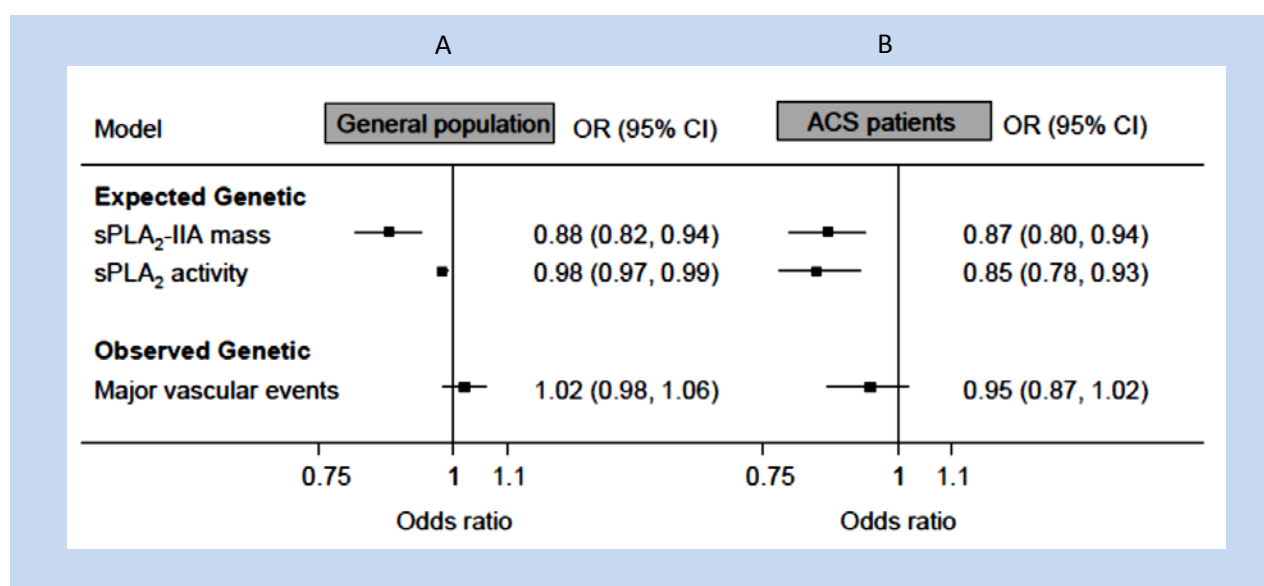


Figure 4.9. Observed and expected genetic effects of rs11573156 on major vascular events assuming unbiased associations of sPLA₂-IIA mass and sPLA₂ activity on fatal/nonfatal myocardial infarction in (A) general population (incident events); (B) ACS patients (recurrent events). The expected genetic effects were estimated twice, first for sPLA₂-IIA mass and then sPLA₂ activity. For each of the expected effects, the effect of the rs11573156 variant on sPLA₂-IIA mass and sPLA₂ activity was used, as well as the effect of sPLA₂-IIA mass and sPLA₂ activity on fatal/nonfatal MI. The observed genetic effects are derived from the per C allele association of rs11573156 with major vascular events (comprising fatal/nonfatal MI or stroke).

4.3. Discussion

The aim of this chapter was to try and accurately classify the causal role of sPLA2-IIA on cardiovascular events using a genetic approach. This process would also allow us to evaluate the worth of sPLA2 inhibition as a specific therapeutic target for decreasing the risk of cardiovascular disease. The *PLA2G2A* functional SNP that I identified in **Chapter 3**, rs11573156, was shown to have a large effect on circulating sPLA2-IIA mass and a significant, albeit more modest, effect on sPLA2 activity, proving it a strong genetic instrument for the evaluation of the role of sPLA2 in disease. No association was found between rs11573156 and prevalent, incident or recurrent myocardial infarction or stroke, which strongly suggests that earlier observation associations between sPLA2-IIA mass and cardiovascular events are due to residual confounding or reverse causality. This is supported by the results seen in Table 45A and 45B, which show that several known cardiovascular risk factors are significantly associated with increasing tertiles of sPLA2 activity and sPLA2-IIA mass. Adjustments for these biomarkers weakens the significance of the effect of sPLA2 on CHD risk, suggesting that sPLA2 may act through one or more of these confounding biomarkers to effect the risk of CHD. The results from the 5 studies originally genotyped for rs11573156 did show a significant association between the SNP and CHD after adjustment, however the loss of this significant association in an much larger meta-analysis leads back to these original results and an increased likelihood that sPLA2 acts through reverse causation or one of the known CHD risk markers that the SNP was shown to be associated with.

The sPLA2 specific inhibitor, varespladib, has been developed in the last several years by Anthera Pharmaceuticals, who hoped to use the drug in conjunction with atorvastatin to decrease the risk of secondary cardiac events in patients with acute coronary syndrome. The phase III study VISTA-16 was set up, due for completion in 2012, which looked to recruit 6,500 ACS patients across Europe and the USA (Nicholls, Cavender et al. 2011). The data has major clinical impact since the VISTA-16 trial (Nicholls, Cavender et al. 2011) (<http://investor.anthera.com/releasedetail.cfm?ReleaseID=656340>) was prematurely halted due to efficacy issues. The 6,500 ACS patients were to be randomised between placebo and a 500mg/day dose of varespladib for 16 weeks. The primary combined end-point was cardiovascular death, nonfatal myocardial infarction, nonfatal stroke or documented unstable angina. Design issues included the relatively brief duration of therapy and the low number of subjects, which may have explained the lack of benefit from the drug. However, this genetic data now suggests that in reality sPLA2-IIA is highly unlikely to be a valid therapeutic target for the prevention of primary or secondary cardiovascular events and this is therefore the likely reason for the lack of benefit seen with varespladib.

Varespladib is known to inhibit 3 forms of sPLA2 enzyme, IIA, V and X. However, the IC_{50} values for the drug show it to be most effective at reducing sPLA2-IIA in mice and human atherosclerotic plaques (IC_{50} 9nM compared to 77nM and 15nM for groups V and X respectively) (Fraser, Hislop et al. 2009). The drug was shown to reduce sPLA2-IIA mass by 78% (Rosenson, Elliott et al. 2011) and thus my SNP of interest appears to be the strongest single candidate that could be used for this type of analysis. The results of our large scale meta-analysis showed null associations between rs11573156 C>G and all CHD/CVD related traits, with the only significant associations were between the SNP and sPLA2-IIA mass/sPLA2 activity. The MR results using rs11573156 suggest that differential sPLA2-IIA levels are not causally associated with CHD and the differential levels of sPLA2-IIA associated with changes in CHD seen in observational studies are more likely as a product of the onset of CHD.

Since the SNP that I identified for use as a genetic instrument was specific to *PLA2G2A*, we saw a very strong and specific association between the SNP and sPLA2-IIA mass and a less strong association with sPLA2 activity, which also includes sPLA2-V and sPLA2-X in its measurement, suggesting that the sPLA2 isoforms, sPLA2-V and X may make significant contributions to sPLA2 activity even with lower measures of sPLA2-IIA. The sPLA2-V and X isoforms and their potential effects are explored in greater detail in **Chapter 5 and 6**.

The fact that the analysis carried out here does not cover the sPLA2-V or sPLA2-X isoforms is a limitation, since we cannot exclude the possibility that there is a causal role for one of these isoforms for cardiovascular events. A further limitation is the smaller number of subjects in the ACS set studies, meaning that there was less power to exclude a small association with myocardial infarction or stroke if it were real, as demonstrated by the small overlap of the expected and observed estimates for sPLA2-IIA mass and sPLA2 activity with CVD in an ACS setting. However, comparing individuals homozygous for the rs11573156 C allele to those homozygous for the G allele resulted in a comparable reduction in sPLA2-IIA mass to the effect of 500mg/day varespladib, and also showed no association with incident or recurrent MVE, suggesting that the lack of efficacy of varespladib for the reduction of recurrent MVE is a true result. A final limitation for this chapter is that there was not data from a common set of participants with measures of sPLA2 and cardiovascular disease, as is often the case with large-scale meta-analyses of Mendelian randomization studies (Minelli, Thompson et al. 2004). However, in this study, we identified a very strong association between rs11573156 and sPLA2-IIA in the subset of studies with measures of sPLA2-IIA, reducing the risk of measurement error in this stage.

As previously acknowledged the genetic instrument I identified is specific for sPLA2-IIA and although meta-analysis was more than sufficiently powered, there were a much smaller proportion of studies examining secondary cardiac events. Therefore my prediction would be that varespladib and similar sPLA2 inhibitors will have no effect on either primary or secondary heart disease, but further studies into the effects on secondary events and the particular contributions of sPLA2-V and X will be necessary.

The independent risk factor status of sPLA2-IIA seen in the early observational and *in vitro* studies was thus not reflected by the studies shown in this thesis, suggesting that sPLA2-IIA is merely a marker for atherosclerosis, perhaps acting through a confounding known cardiovascular risk trait as suggested by the results in Table 45A-B, or through reverse causation, but it does not appear to be directly causal.

4.4. Conclusion

This collaboration has yielded a large-scale Mendelian randomisation study that does not support the hypothesis that sPLA2-IIA mass is causally linked to cardiovascular events or that it is a viable drug-target for cardiovascular prevention. The parallel between these genetic findings and the lack of efficacy of the sPLA2 inhibitor, varespladib, seen in the VISTA-16 trial supports the use of this genetic approach earlier in drug development, as this may help to prioritise the best drug targets to follow in human randomised trials and this in turn could save a lot of time and money that is often wasted due to late-stage drug failure.

CHAPTER 5

The Potential Contribution of sPLA2-V to sPLA2 activity and Risk of CHD

5.1. Introduction

5.1.1. Background: sPLA2 Isoenzyme Analysis

Past research has shown that the sPLA2 group of enzymes are classified together due to a range of structural similarities as described in **Chapter 1**. As well as their common ability to hydrolyse the sn-2 ester bonds of phospholipids the sPLA2 enzymes generally have comparable low molecular weights (~14kDa), similar secretory mechanisms, a histidine/calcium dependency and are generally rich in disulphide bonds (Six and Dennis 2000; Masuda, Murakami et al. 2005; Lambeau and Gelb 2008). However, the enzymes are defined individually by the subtle differences in each protein structure that lead to distinctions in the mechanism and efficiency of phospholipid hydrolysis for each enzyme. These structural changes often relate to differences in the overall charge of each enzyme (Lambeau and Gelb 2008). It has also been proved by immunohistochemical staining that each enzyme has a different expression profile in different tissues, suggesting alternative roles for each isoenzyme (Masuda, Murakami et al. 2005; Lambeau and Gelb 2008). These differences highlight the importance of trying to interpret the contribution of each individual isoenzyme to sPLA2 activity and ultimately to diseases such as heart disease. The results presented in **Chapter 4** have shown that sPLA2-IIA is highly unlikely to contribute to the onset of CHD, but since the MR analysis that was carried out in **Chapter 4** did not give us a true indication of the potential causal contributions of other isoenzymes, such as sPLA2-V, to CHD it was important to establish whether sPLA2-V could potentially have an independent causal association with CHD (**5.1.2- 5.1.5**) and, if so, to begin the process of identifying potentially functional SNPs in these genes to be used as instruments for MR as with rs11573156 for *PLA2G2A* (**Chapter 3** and **Chapter 4**). Further to this, examining the potential contribution of each *PLA2* gene expression pattern to sPLA2 activity was important in order to present a clearer picture of how these enzymes might interact and what the potential significance of sPLA2 inhibition could be depending on the contribution of each protein.

5.1.2. Comparing sPLA2-V and sPLA2-IIA

As I have previously discussed, sPLA2-IIA and sPLA2-V share many common features and the genes *PLA2G2A* and *PLA2G5* are transcribed in opposite directions in a cluster on chromosome 1, only 90kb apart (**Chapter 1**, Figure 1.11). I hoped to use genetic analysis of SNPs in the *PLA2G5* gene to help to explain the differing functional roles played by sPLA2-V and sPLA2-IIA in sPLA2 activity and in CHD. (Tischfield, Xia et al. 1996; Wootton, Arora et al. 2007).

A deficiency in any individual sPLA2 enzyme can affect total sPLA2 activity in different ways and often in a tissue dependent manner (Murakami, Taketomi et al. 2011). This may indicate distinct functional roles for sPLA2-IIA and sPLA2-V and since these potential enzyme deficiencies could affect activity to different magnitudes, in different tissues depending on which isoenzyme is depleted, it is reasonable to assume that sPLA2-V could have different effects in diseases such as atherosclerosis compared to sPLA2-IIA. However there is also some evidence that suggests that a deficiency in one could be partially compensated for by the presence of the other, suggesting that these enzymes may work reciprocally (Sawada, Murakami et al. 1999) and one of the aims of this thesis was to try and identify what the distinct roles for each enzyme might be.

5.1.3. The Potential Contribution of sPLA2-V to sPLA2 Activity

Since the fluorometric assay used to measure sPLA2 activity is inclusive of three sPLA2 isoforms; groups IIA, V and X, research has so far been unable to successfully identify the individual contributions of each of these enzymes to sPLA2 activity and this may be important when determining the contributions of each isoform to diseases such as CHD.

As documented previously, sPLA2-V has around 20 times higher potency towards common phospholipids compared to sPLA2-IIA due to their different structural elements and charges (Lambeau and Gelb 2008). These differences mean that while both group IIA and group V enzymatic properties include proteoglycan binding and the release of arachidonate through a Heparan Sulphate proteoglycan (HSPG) shuttling pathway, only group V has the ability to additionally release arachidonate through the external plasma membrane pathway (Figure 1.12), albeit to a lesser extent than through the HSPG pathway (Murakami and Kudo 2003). The increased potency of sPLA2-V compared to sPLA2-IIA could suggest that group V may contribute more to overall sPLA2 activity, however sPLA2-IIA levels are significantly higher than the other sPLA2 isoforms, which may be more important than potency in a natural physiological state (Lambeau and Gelb 2008). When studying the contributions of each enzyme to sPLA2 activity it is important to take into account where each enzyme is

expressed, since differential expression patterns exist for each isoform in each tissue dependent on the circumstances (e.g. healthy versus diseased tissue) (**Chapter 1.5**).

5.1.4. Investigating the Contribution of sPLA2-V to CHD Risk

Chapter **1.5.7** explores the potential role of sPLA2-V in CHD risk and discusses evidence from *in vitro* and *in vivo* studies to support these claims. Chapter **1.7.2** goes on to identify a study (Wootton, Arora et al. 2007) that has previously isolated 7 tagging-SNPs that span the gene (**1.7.2** Table 1.3 and Figure 1.13). A further SNP of interest was identified by analysis carried out by Dr Lasse Folkersen in the ASAP study. In this instance an eQTL with the strongest differential association with *PLA2G5* mRNA was identified as the variant rs525380 (Figure 5.1). In this chapter I describe some of the expression patterns of *PLA2G5* and related tagging-SNPs to attempt to provide useful insights into the tissue distribution of *PLA2G5* expression and to use genetics to examine any associations between *PLA2G5* variations, sPLA2 activity and risk of CHD/CHD traits in a small meta-analysis setting.

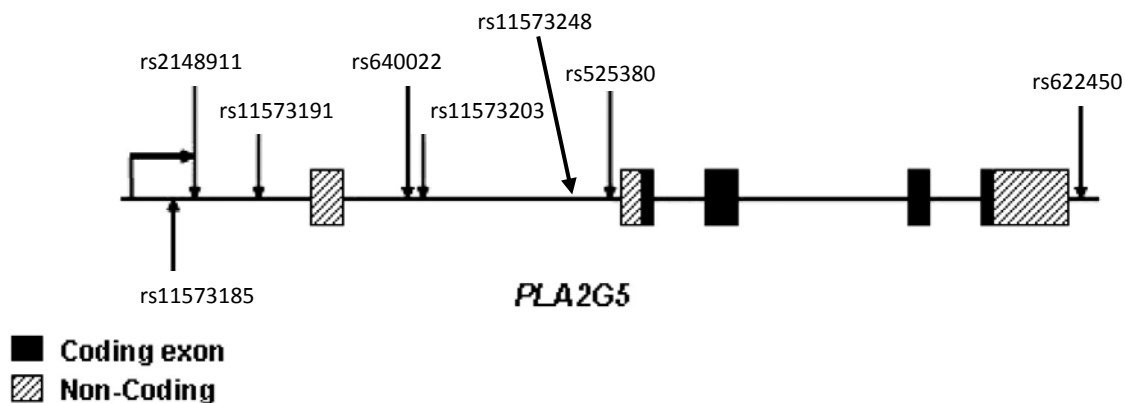


Figure 5.1. A schematic of the *PLA2G5* gene and the 7 tSNPs originally identified (Wootton, Arora et al. 2007) as well as the SNP identified by eQTL data as significantly differentially associated with *PLA2G5* mRNA, rs525380 A>C. Schematic reprinted from (Wootton, Arora et al. 2007). The region spanned here is from the upstream 5'UTR SNP rs11573185 to downstream 3'UTR SNP rs622450, since these were the most extremely located previously identified tSNPs; Chr1: 20,395,401-20,419,340.

5.1.5 SNP Selection

The SNPs selected from the study by Wootton et al. for further analysis included; rs11573185 C>A, rs11573191 A>G, rs11573203 G indel and rs11573248 TA indel. These 4 SNPs had been shown by adjusted univariate analysis in UDACS to be significantly associated with cholesterol (a known CHD biomarker) and sPLA2-IIA levels (Table 5.1) (Wootton, Arora et al. 2007). The SNP rs11573185 was the only SNP from this group available on the ASAP array. Since this SNP was also shown to be associated with sPLA2-IIA levels, it was taken forward for further analysis in the ASAP study. The SNPs rs11573248 and rs11573203 may warrant further genotyping since there is very little known genotyping data for these SNPs.

SNP/Genotype:	N	log sPLA2-IIA Mass (ng/ml)	P value unadjusted	P value adjusted*	P value adjusted**	LDL-C (mmol/l)***	P value unadjusted	P value adjusted*	P value adjusted**
rs11573185 11	105	2.76	0.10	0.06	0.04	2.85	0.08	0.01	0.02
rs11573185 12	244	3.22				2.7			
rs11573185 22	155	3.32				2.58			
rs11573191 11	347	3.36	0.004	0.004	0.002	2.62	0.02	0.01	0.02
rs11573191 12	146	2.90				2.88			
rs11573191 22	14	1.93				2.79			
rs11573203 11	261	3.39	0.006	0.008	0.007	2.58	0.02	0.007	0.01
rs11573203 12	190	3.05				2.80			
rs11573203 22	39	2.33				2.92			
rs11573248 11	210	3.01	0.08	0.05	0.02	2.84	0.007	0.0003	0.0005
rs11573248 12	227	3.17				2.64			
rs11573248 22	69	3.78				2.46			

* = adjusted for age, gender, CHD status and statin use

**= adjusted for age, gender, CHD status, CRP and statin use

***= square root transformed

Table 5.1. Adjusted and unadjusted univariate analysis for significant associations found with four *PLA2G5* tSNPs identified in the 2007 study by Wootton et al. with sPLA2-IIA mass and the known CHD biomarker, LDL-cholesterol. The analysis is based on genotyping data for each SNP in the UDAC Study (previously defined in **2.5**)

Bioinformatics of tSNPs

A detailed bioinformatics analysis was carried out for both rs11573185 and rs525380. Considering the other three tSNPs identified along with rs11573185, analysis from the UCSC Genome Browser showed that only rs11573248 could potentially be functional since it exists in an area of DNA hypersensitivity and in a potential transcription factor binding region (Figure 5.6). Further genotyping analysis of this SNP may be useful to investigate the relationship between sPLA2-V and sPLA2-IIA as this was also the tSNP with the most significant differential association to sPLA2-IIA levels (Wootton, Arora et al. 2007).

Additionally the SNP rs525380 (Figure 5.1) was identified as a potential SNP of interest from initial ASAP study analysis (Figure 5.7) since it showed the highest association with differential *PLA2G5* mRNA in the aortic adventitia. Based on these results rs525380 was analysed further and genotyping has been completed in EPIC-Norfolk with additional genotyping data from WTCCC. Analysis in this study was carried out by Dr Reecha Sofat. Due to their inclusion in the ASAP arrays, rs11573185 and rs525380 are considered the 'lead' SNPs in this results chapter.

5.2. Results

5.2.1. Bioinformatics Analysis

5.2.1.1. Bioinformatics Analysis of *PLA2G5*

According to the Ensembl Genome Browser and NCBI databases there are 8 known transcripts of the gene but only one of these is protein coding. The protein coding transcript is made up of 5 exons in total, 4 coding and 1 non-coding (Figure 1.13 and Figure 5.1). This transcript codes for a 138 amino acid protein. There is supporting evidence for this transcript being protein coding including protein evidence, EST evidence and cDNA copy evidence. The additional 7 transcripts are identified by Ensembl as non-coding since they do not contain an open reading frame. Supporting evidence for these transcripts includes cDNA and EST but not protein evidence.

5.2.1.2. Bioinformatics Analysis of the Lead tSNP rs11573185 C>A

Data from NCBI and Ensembl verified that rs11573185 C>A was positioned at chromosome 1:20,395,401 which is in the region immediately upstream of the 5'UTR of *PLA2G5* (Figure 5.2). The genotyping data for Europeans available from HapMap release #28 and the 1000 Genomes Project showed that the rare A allele occurs at a frequency of 0.47. Since this SNP is a tagging-SNP it is in strong LD (D') with several other SNPs; Figure

5.3 is an LD heat plot representing *PLA2G5* and HapMap release #28 genotyped SNPs, including rs11573185. The higher the intensity of red, the higher the D' LD score is between 2 SNPs. Where acknowledged, the R^2 values appear as numbers in each block. As is shown on the LD map, rs11573185 exists in a region of high D' with its surrounding SNPs from Chr1: 20,394,713-20,395,868 (Figure 5.3). However, as shown in Figure 5.4, taken from the SNAP pairwise LD plotting tool using the 1000 Genomes Pilot 1 data, this SNP does not have significant R^2 ($R^2 < 0.7$) with any known SNPs across *PLA2G2A* and *PLA2G5* in this database. This is probably due to the fact that the 1000 Genomes Pilot 1 data is a subset of the HapMap CEU data and therefore includes less individuals from the same sample and less power to show LD between SNPs. The CEU population for both samples (1000 Genomes Pilot 1 and HapMap release #28) are based on Utah residents with Northern and Western European ancestry from the CEPH collection. The 1000 Genomes population is based on 60 of the HapMap genotyped individuals and the HapMap population is based on 226 genotyped individuals.

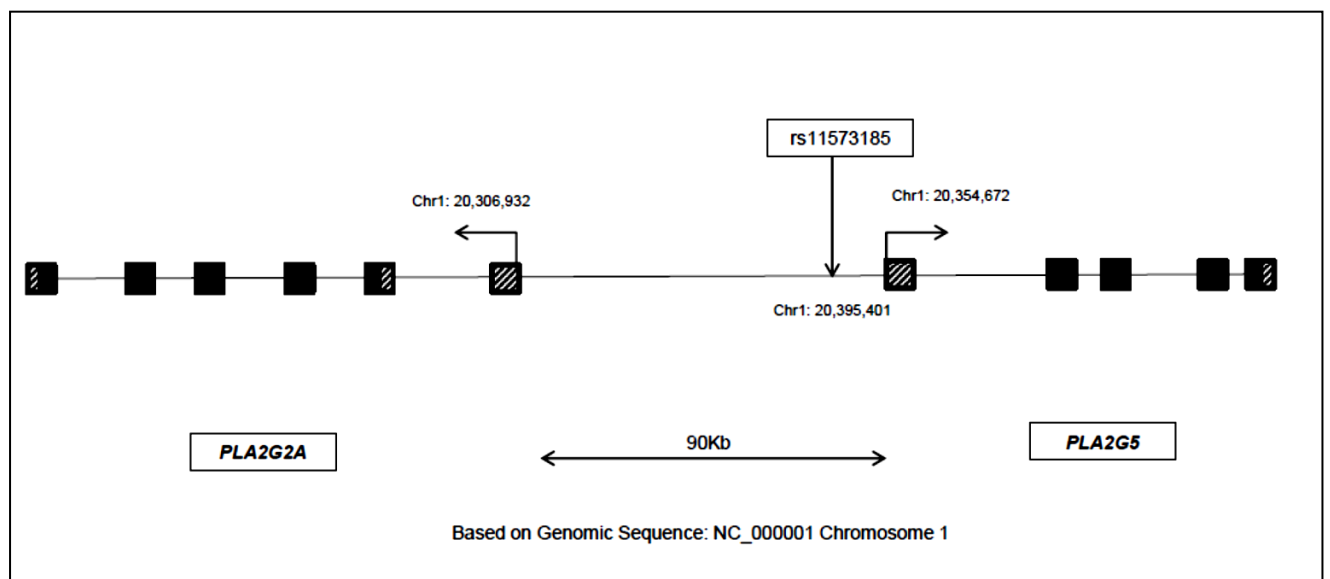


Figure 5.2. Shows the location of rs11573185 upstream of the 5'UTR of *PLA2G5* within the 90Kb shared region between *PLA2G2A* and *PLA2G5*. Hatched boxes represent non-coding exons. Filled boxes represent coding exons. The arrows on exons indicate the direction of transcription for each gene (reverse transcription for *PLA2G2A* and forward transcription for *PLA2G5*). The figure is based on the genomic sequence NC_000001 on chromosome 1 and data from the Ensembl Genome Browser.

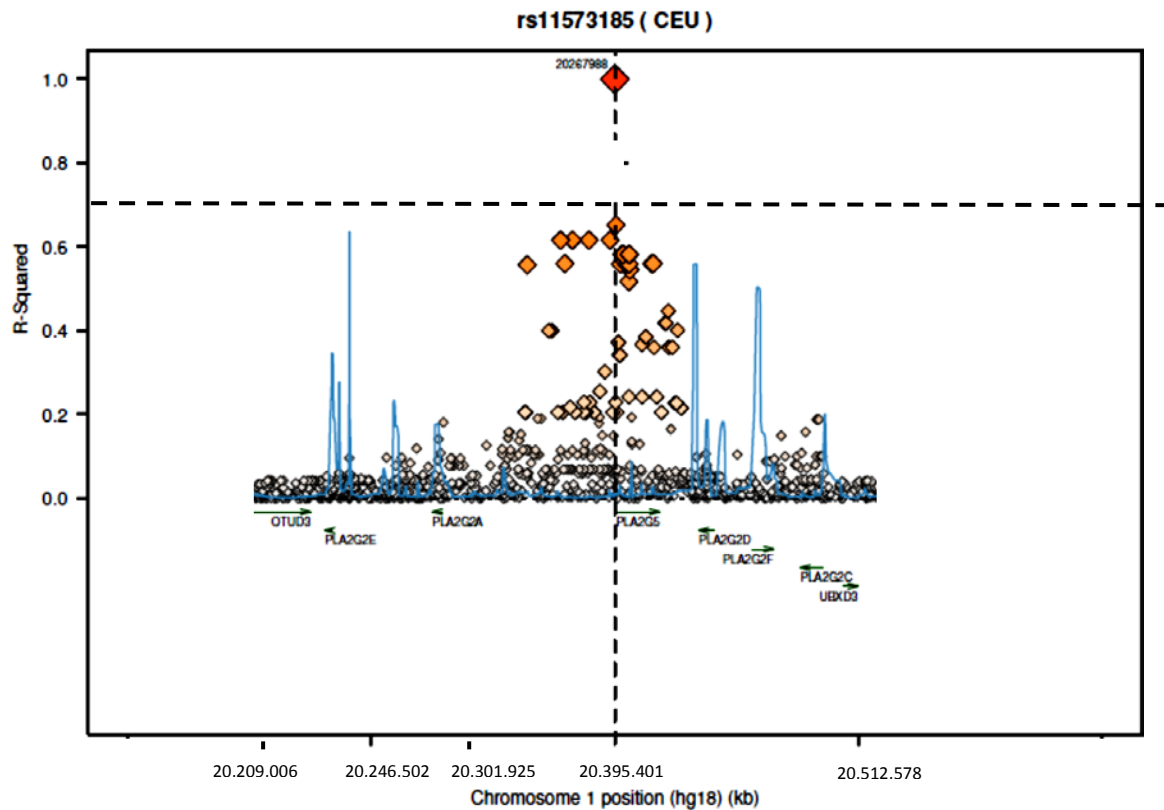


Figure 5.4. An LD Manhattan style plot based on the 1000 Genomes CEU pilot 1 data available through the SNAP Pairwise LD program. The SNAP plot tool was used to assess the LD of rs11573185 with all known SNPs in the region of 120Kb upstream (as far as the *OTUD3* gene) and 185Kb downstream (as far as the *UBXN10* gene) of the SNP, which covers the whole *PLA2* gene cluster. The location of rs11573185 is represented by the vertical dashed line. Genotyped and imputed SNPs are represented by a diamond shape. The R^2 significance value is set at 0.7 and represented by a horizontal dashed line. The SNPs in LD above 0.7 are indicated above this line. The deeper the shade of red and the larger the size of each diamond, the stronger the LD with rs11573185. The blue peaks on the plot represent recombination hotspots and the green arrows represent different genes. The measure of LD calculated as R^2 is indicated by Y axis values.

The UCSC Genome Browser identifies no potential transcription factor binding region in the sequence immediately around rs11573185 (Figure 5.6). Additionally, it does not appear near a known splicing region, or in a DNA hypersensitive region. This data suggests that if any associations are seen between rs11573185 and either sPLA2, CHD or CHD risk factor measures, it is more likely to be due to a SNP that rs11573185 is tagging as opposed to directly due to this SNP. This may be a previously defined SNP in LD with rs11573185 or it is possible that a true functional SNP has not been genotyped yet, or is in a region unrelated to the *PLA2* cluster.

5.2.1.3. Bioinformatics Analysis of rs525380 A>C

Data from NCBI and Ensembl verifies that rs525380 A>C is positioned at chromosome 1:20,409,374, which is positioned 1,939bp upstream of the intron 1/exon 2 boundary of *PLA2G5*. The genotyping data in Europeans available from HapMap and the 1000 Genomes Project puts the minor allele frequency at 0.44-0.49 (A). This SNP, was identified as significantly associated with *PLA2G5* mRNA expression in the ASAP study (Figure 5.7). According to the data from both HapMap European and 1000 genomes European genotyping data this SNP tags rs576352 (*PLA2G5* intron 1 Chr1: 20,397,157) rs513246 (*PLA2G5* intron 1, Chr1:20,398,043) and rs656110 (*PLA2G5* intron 1, Chr1:20,397,240). An LD plot with rs525380 as the SNP of interest shows a significant threshold for R^2 at 0.7 and is shown in Figure 5.5, taken from the SNAP pairwise LD plotting tool. This plot covers *PLA2G2A*, *PLA2G5* and other *PLA2* cluster SNPs. The CEU population for both samples (1000 Genomes Pilot 1 and HapMap release #28) are based on Utah residents with Northern and Western European ancestry from the CEPH collection. The 1000 Genomes population is based on 60 of the HapMap genotyped individuals and the HapMap population is based on 226 genotyped individuals.

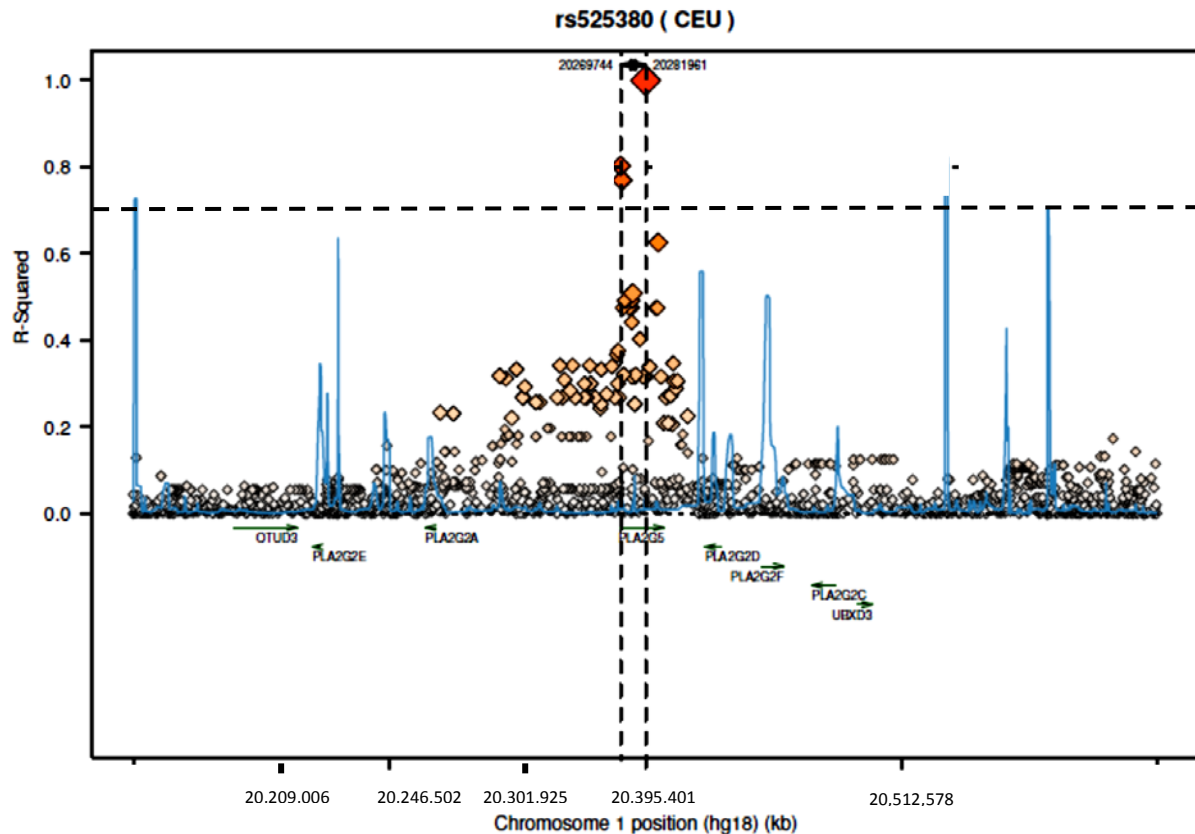


Figure 5.5. An LD Manhattan style plot based on the 1000 Genomes CEU pilot 1 data available through the SNAP Pairwise LD program. The SNAP plot tool was used to assess the LD of rs525380 with all known SNPs in the region of 500Kb upstream and 500Kb downstream of the SNP. The region that includes SNPs in LD <0.7 with rs525380 are represented by the vertical dashed lines. Genotyped and imputed SNPs are represented by a diamond shape. The R^2 significance value is set at 0.7 and represented by a horizontal dashed line. The SNPs in LD above 0.7 are indicated above this line. The deeper the shade of red and the larger the size of each diamond, the higher the strength of LD with rs525380. The blue peaks on the plot represent recombination hotspots and the green arrows represent different genes.

When rs525380 was input to the UCSC Genome Browser no potential transcription factor binding regions could be identified in the sequence immediately around the SNP (Figure 5.6). Additionally, it does not appear near a known splicing region, or in a DNA hypersensitive region. This data suggests that if any associations are seen between rs525380 and either sPLA2, CHD or CHD risk factor measures, it is more likely to be due to a SNP that rs525380 is tagging as opposed to directly due to this SNP.

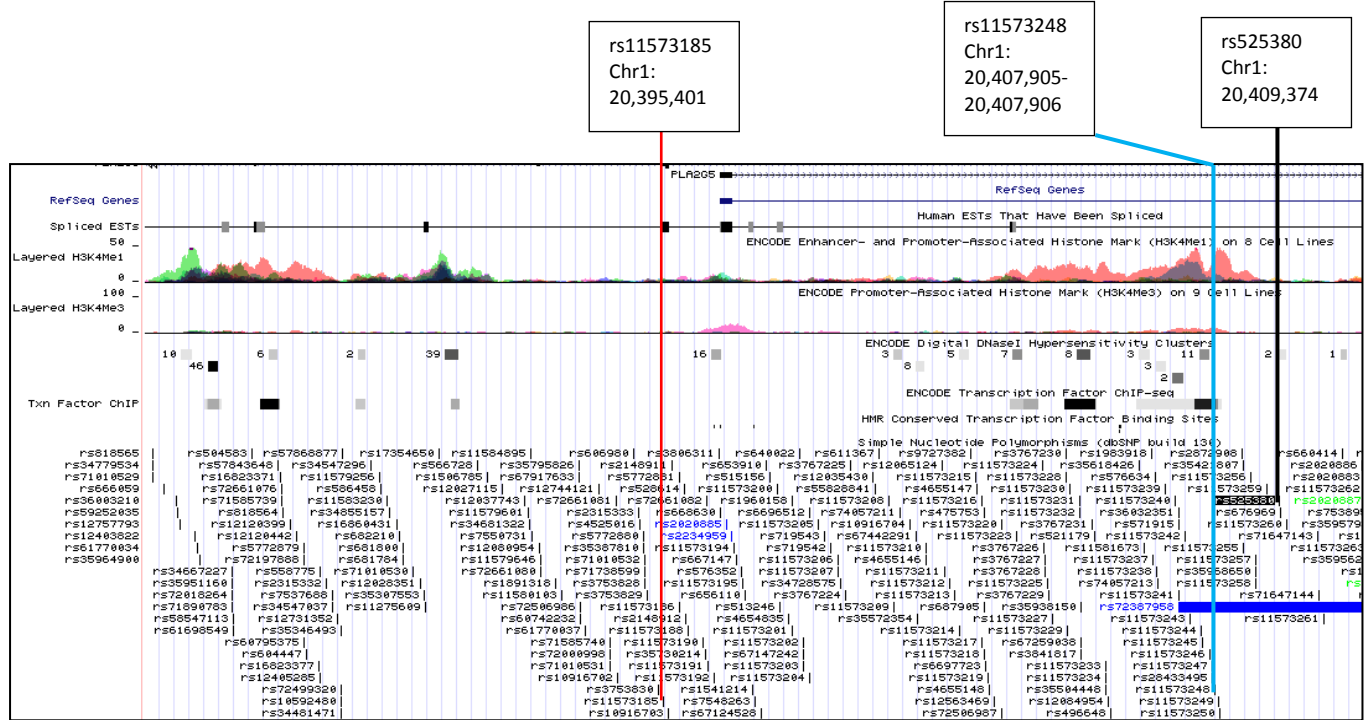


Figure 5.6. Has been reprinted using data from the UCSC Genome Browser for *PLA2G2A*. The gene is shown on the forward strand. The protein coding transcript is shown based on data from RefSeq, UniProt, GenBank, CCDS and Comparative Genomics. Spliced ESTs represents Human expressed sequence tags (short strands of cDNA used to identify gene transcripts), which have been used to identify potential splice sites within the gene. These are shown in black. Layered H3K4Me1 represents data from Encode ChIP sequence assays to show the level of predicted histone 3 and 4 methylation/acetylation. Each colour represents a different cell type in which the assay was carried out. Layered H3K4Me3 shows DNase hypersensitivity regions are shown in the same cell lines. The ENCODE Transcription Factor ChIP-seq data shows the prediction of transcription factors binding to the gene. These are labelled accordingly. The intensity of colour indicated the likelihood of true binding. Finally all the SNPs covering the region are indicated by rs number at the bottom. The SNPs of interest, rs11573185 and rs525380, are indicated by the vertical red line and the vertical black line respectively and the SNP rs11573248 is indicated by the vertical aqua line (UCSC Genome Browser) (<http://genome.ucsc.edu/cgi-bin/hgTracks?position=chr1:202802420280743&hgside=282534055&snp130=pack&hgFind.matches=rs11573248>) (<http://genome.ucsc.edu/cgi-bin/hgTracks>).

5.2.2. *PLA2G5* mRNA Expression in the ASAP Study

PLA2G5 mRNA expression across all the tissues included in the ASAP study (mammary artery, liver, aorta media, aorta adventitia and heart) was characterised by the analysis of all 123 *PLA2G5* SNPs covered by the array, covering a large portion of the gene. The Manhattan plot style Figure 5.6 indicates which SNPs were associated with the strongest differential mRNA expression and in which tissues by colour coding. The data showed that the strongest expression of *PLA2G5* mRNA was most significantly associated with SNPs in the aortic adventitia (aqua) and liver (blue). The SNP rs525380 was the SNP with the single strongest significant association with differential *PLA2G5* mRNA levels as indicated in Figure 5.7 as the highest point on the Manhattan plot. This association is measured in the aorta adventitia. However, as shown previously, across all 5 tissues in the ASAP study the strongest association between rs525380 and *PLA2G5* mRNA levels is

comparatively lower than the strongest SNP association with *PLA2G2A* mRNA (as shown in Figure 4.4).

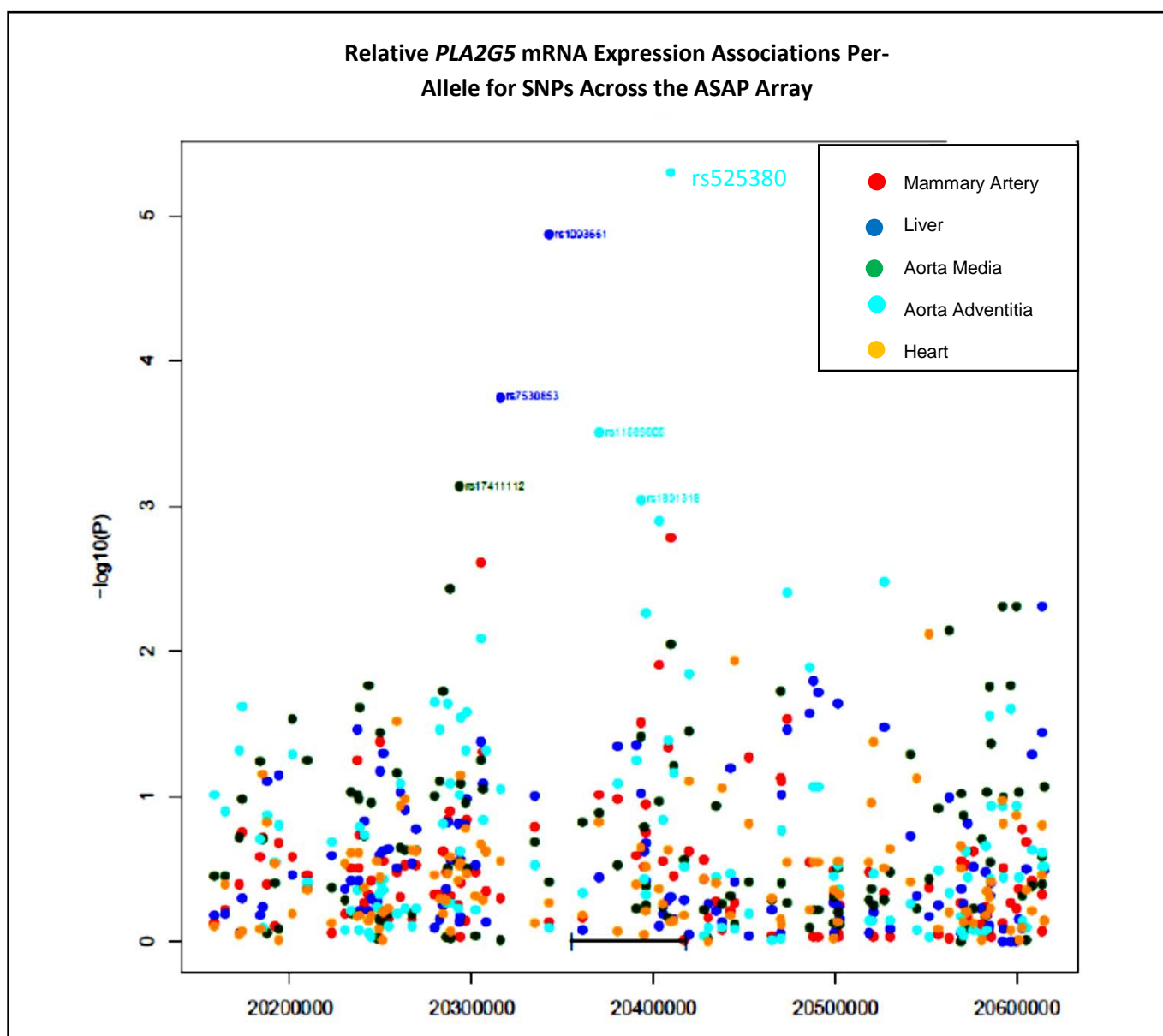


Figure 5.7. A Manhattan plot style figure to show the expression of *PLA2G5* mRNA associated with SNPs across the tissues included in the ASAP study. The strongest single association is seen with the SNP rs525380 A>C in the aorta adventitia. The black horizontal line above the scale represents the position of *PLA2G5*.

Both rs11573185 and rs525380 were analysed using the arrays of the ASAP study. Analysis was performed to assess the levels of both these SNPs across the tissues included in the study; mammary artery, aorta media, aorta adventitia, liver and heart, thus allowing us to identify if there is likely differential expression between the rs11573185 or rs525380

variants in cardiovascular related tissues such as the aortic adventitia, aortic media, heart and liver. The results showed that the SNP rs11573185 was associated with the highest levels *PLA2G5* mRNA most highly in the aortic adventitia, but was associated with the least mRNA expression in the aorta media and in heart tissue (Figure 5.8). Further analysis showed there to be no significant per allele difference in mRNA expression across the aortic adventitia for rs11573185 ($P=0.08$), suggesting that this SNP is not functional regarding *PLA2G5* mRNA expression in this tissue (Figure 5.9). Regarding rs525380, Figure 5.10 shows that the most significant association of differential expression of *PLA2G5* mRNA with rs525380 is in the mammary artery, the aortic media and the aortic adventitia. Figures 5.11A, 5.11B and 5.11C show the per allele effect of rs525380 on mRNA expression of *PLA2G5* in these tissues with the most significant association seen with the aorta adventitia (5.05×10^{-6}) (Figure 5.10).

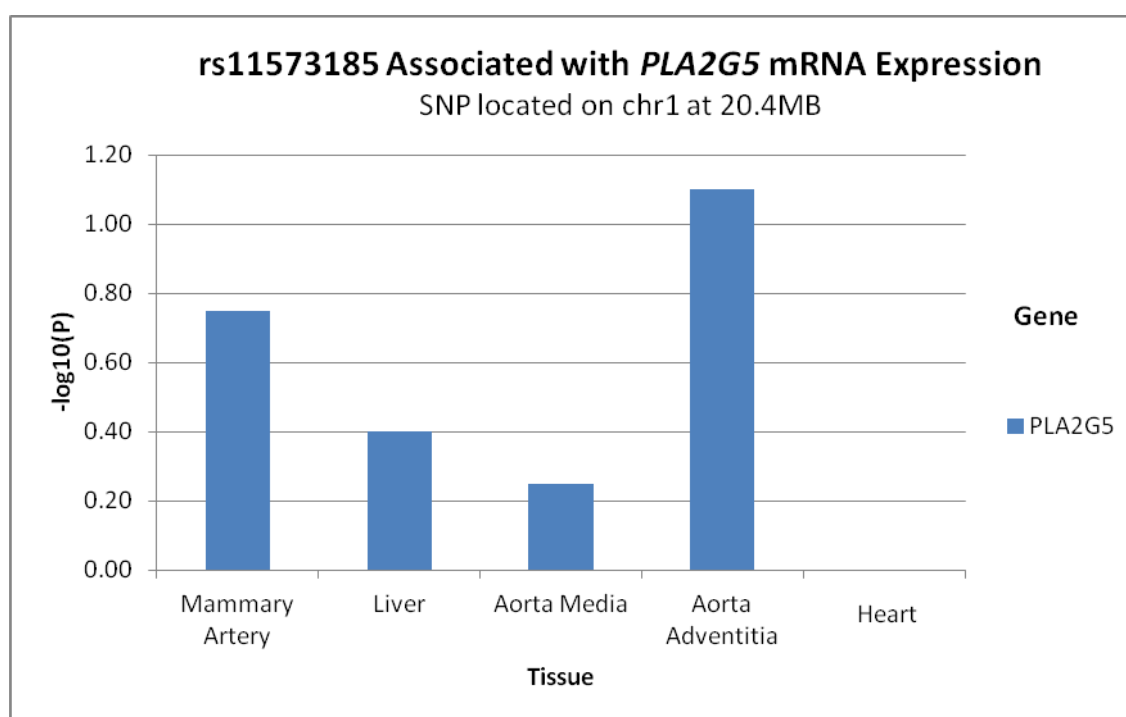


Figure 5.8. Differential *PLA2G5* mRNA expression associated with rs11573185 C>A across all ASAP measured tissues in a per allele additive model. Differential expression was most significantly associated with this SNP in aorta adventitia and least significantly associated in the heart.

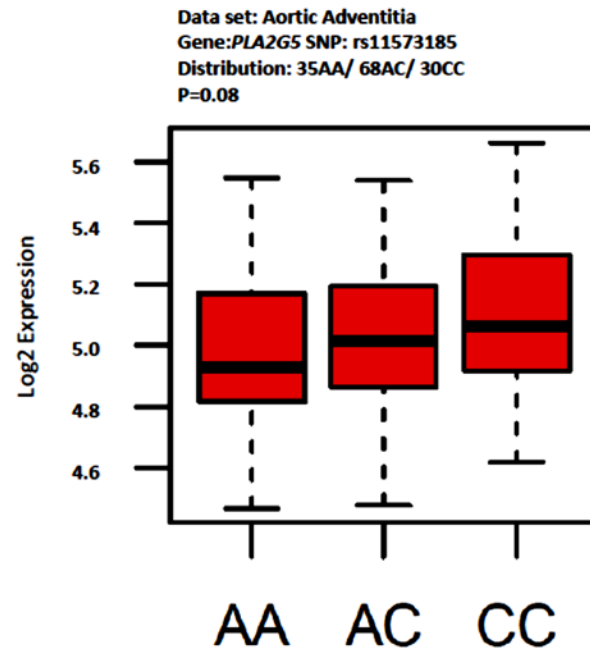


Figure 5.9. The per C allele effect of rs11573185 C>A on *PLA2G5* mRNA expression across the aorta adventitia based on measurements from the ASAP study. The figure shows the effect of each additional C allele increases mRNA expression in this tissue.

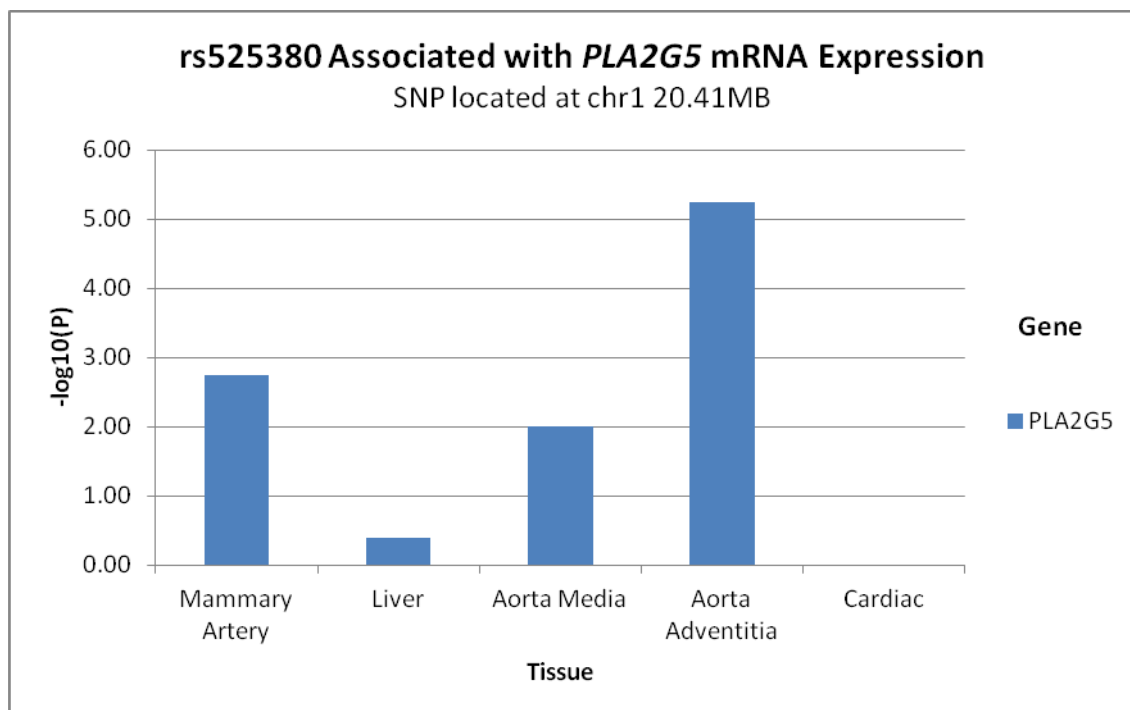


Figure 5.10. Shows the association of rs525380 A>C with differential mRNA expression levels of *PLA2G5* in a per allele additive model. Differential expression is most significantly associated with this SNP in aorta adventitia and least significantly associated in heart.

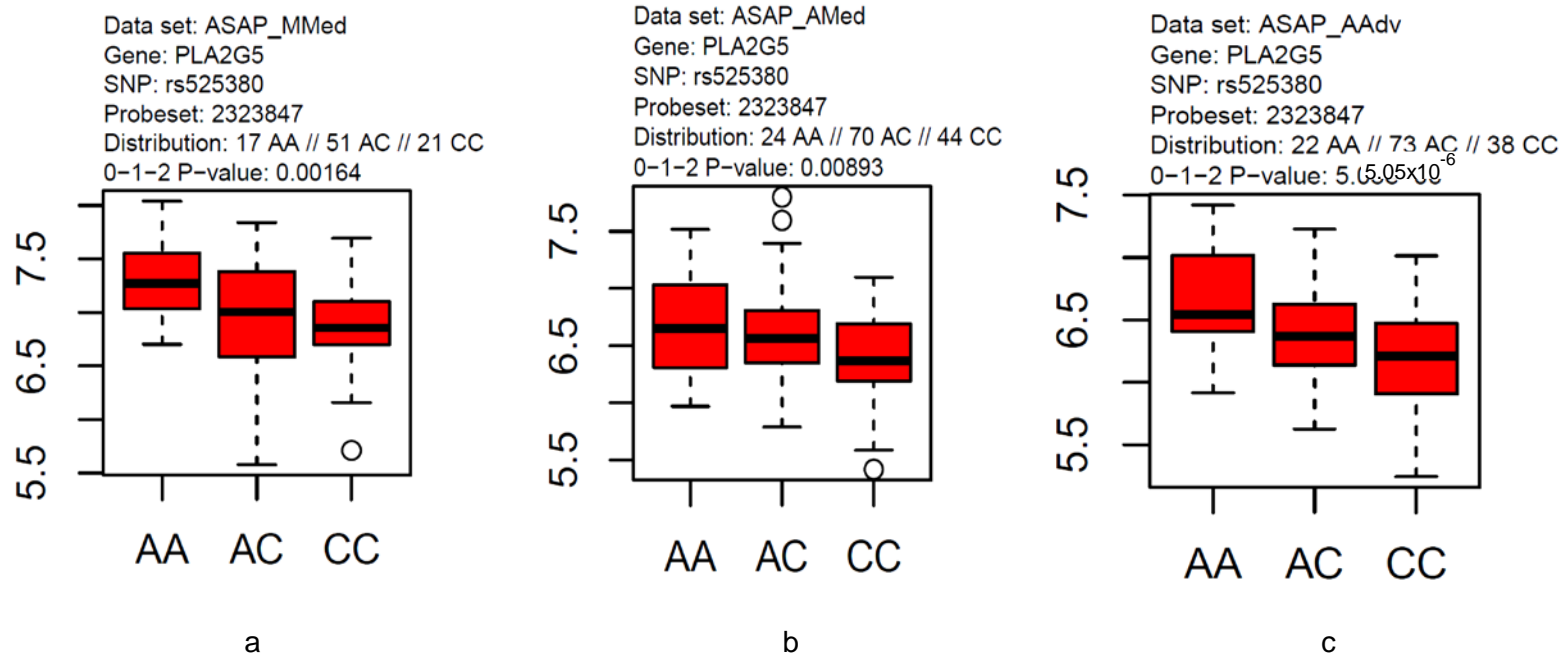


Figure 5.11a Results from the ASAP study showing the rs525380 A>C differential expression association with *PLA2G5* mRNA in the mammary artery. Per C allele there is a reduction in mRNA expression. Figure 5.11b Results from the ASAP study showing the rs525380 A>C differential expression association with *PLA2G5* mRNA in the aorta media. Per C allele there is a reduction in mRNA expression. Figure 5.11c Results from the ASAP study showing the rs525380 differential expression association with *PLA2G5* mRNA in the aorta adventitia. Per C allele there is a reduction in mRNA expression.

5.2.3. Association of *PLA2G5* SNPs with *PLA2G2A* mRNA Levels.

The original data for rs11573185 (Wootton, Arora et al. 2007) shown in Table 5.1 and results from my meta-analysis (Figure 5.18 and Figure 5.21) suggest that the *PLA2G5* SNPs rs11573185 and rs525380 (among others), are significantly associated with sPLA2-IIA levels. I asked Dr Folkersen for data relating to the relative differential expression of *PLA2G2A* mRNA associated with either of these SNPs for me to analyse in order to prove these associations and discount confounding or lack of assay specificity as a cause for these outcomes (Figure 5.12a and 5.12b).

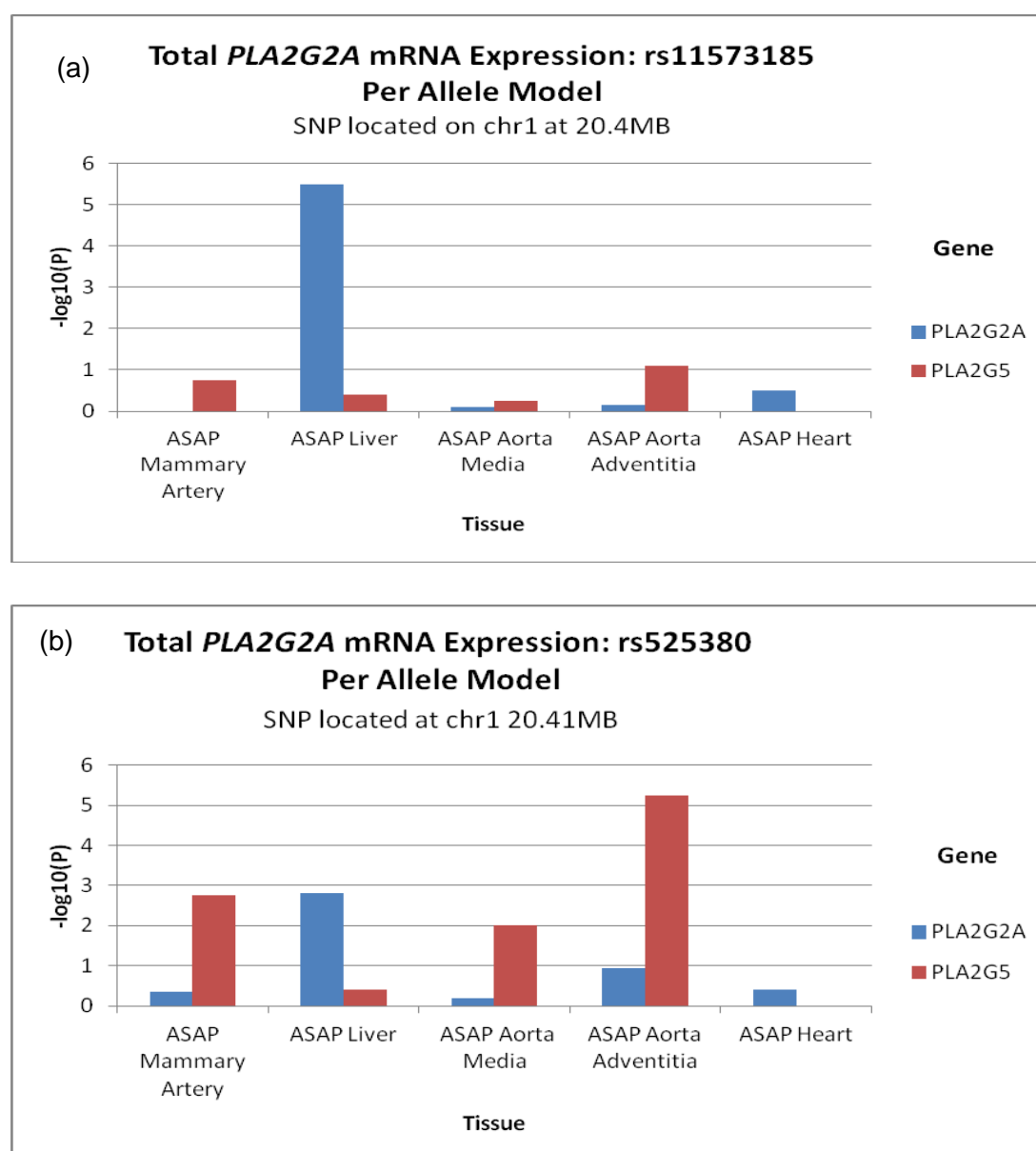


Figure 5.12 a and b show the relative expression of *PLA2G2A* mRNA and *PLA2G5* mRNA with the *PLA2G5* SNPs; rs11573185 (a) and rs525380 (b) across the 5 tissues of the ASAP study; mammary artery, liver, aorta media, aorta adventitia and heart. Both SNPs appear to be associated with *PLA2G2A* mRNA suggesting some cross regulation between *PLA2G2A* and *PLA2G5*.

I then analysed further data provided Dr Folkersen (Figures 5.13a and 5.13b) to confirm that both rs11573185 and rs525380 had significant per-allele associations with *PLA2G2A* mRNA in the liver, suggesting perhaps that *PLA2G2A* and *PLA2G5* are co-regulated. A per C allele increase in *PLA2G2A* mRNA is seen with rs525380 A>C and a per C allele decrease in *PLA2G2A* mRNA is seen with rs11573185 A>C.

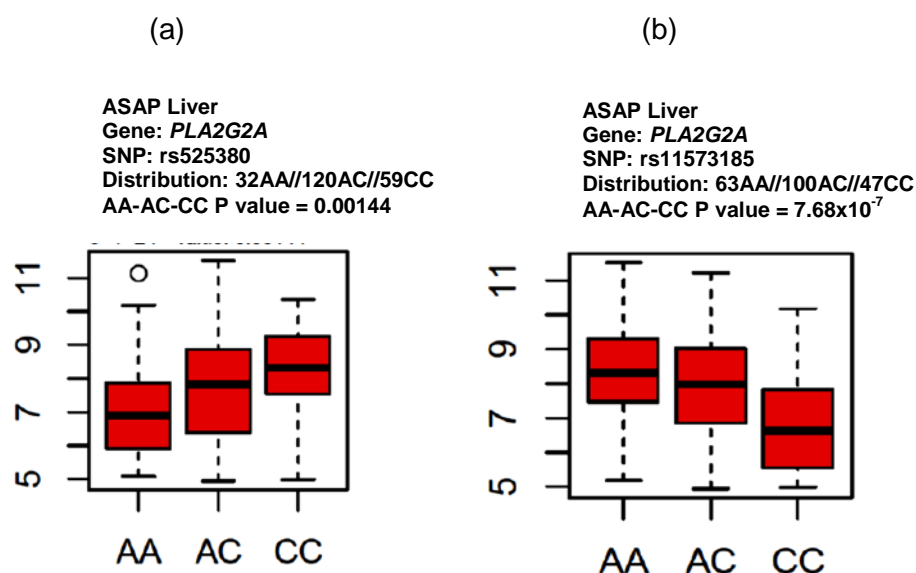


Figure 5.13. Shows the per allele differential *PLA2G2A* mRNA expression patterns associated with (a) rs11573185 A>C and (b) rs525380 A>C in the liver. Both *PLA2G5* SNPs appear to be significantly associated with *PLA2G2A* suggesting some shared regulation between these two genes.

5.2.4. Combined Meta-Analysis Data for rs11573185

Following the identification of four tSNPs of interest in the UDAC study (Wootton, Arora et al. 2007), further genotyping analysis was carried out for each SNP in several additional studies (Table 5.2). No significant associations remained between total cholesterol and any of the selected SNPs and although there was a borderline significant association seen between LDL-cholesterol and rs11573203, this measure was only across 3 studies and the association appears to be driven solely by the UDACS cohort, thus suggesting that this is an anomalous result. None of the selected SNPs showed significant associations with sPLA2 activity or with total CHD in these additional studies. Table 5.2 gives an overview of relevant SNP/biomarker associations across the studies indicated and previously described in Table 2.6. A further in depth analysis for rs11573185 is presented in more detail in Figures 5.14-5.19. All analysis was adjusted for age and gender. Since the four tSNPs shown here were shown by Wootton et al. (Wootton, Arora et al. 2007) to have significant associations

with both sPLA2-IIA mass and cholesterol, which is a known CHD biomarker, the results in Table 5.2 are relative to sPLA2-IIA mass, sPLA2 activity and both LDL and HDL cholesterol where these measures were possible. The association of each SNP with CHD across all genotyped studies is also shown in order to assess the potential likelihood that sPLA2-V is causal for CHD. The data from these studies should be considered as a preliminary meta-analysis. In order to confirm the data as accurate, further genotyping should be carried out to increase the power of the meta-analysis in both the general population and ACS study settings.

SNP:	Biomarker/Trait:	Studies Genotyped*:	P Value:	Notes:	HWE**
rs11573185	LDL-cholesterol	NPHSII (95), EPIC (98), EAS (94), CYPRUS (97), BWHHS, WHII, EPIC-NL	0.41		✓
	HDL Cholesterol	NPHSII (95), EPIC (98), EAS (94), CYPRUS (97), BWHHS, WHII, EPIC-NL	0.50		✓
	sPLA2-IIA Levels	EPIC (98), UDACS (98)	3.35x10 ⁻³⁸		✓
	sPLA2 Activity	EPIC (98)	0.29		✓
	CHD	NPHSII (95), EPIC (98), TPT (96), EAS (94), CYPRUS (97), BWHHS, WHII, EPIC-NL	0.06		✓
rs11573248	LDL-cholesterol	NPHSII (98), EPIC (95), EAS (98), CYPRUS (98), BWHHS	0.82		✓
	HDL Cholesterol	NPHSII (98), EPIC (95), EAS (98), CYPRUS (98), BWHHS	0.18		✓
	sPLA2-IIA Levels	EPIC (95), UDACS (98)	1.36x10 ⁻⁴⁰		✓
	sPLA2 Activity	EPIC (95)	0.15		✓
	CHD	NPHSII (98), EPIC (95), TPT (95), IMPROVE, EAS (98), CYPRUS (98)	0.11		✓
rs11573191	LDL-cholesterol	NPHSII (98), EPIC (97), UDACS (99), GRACE	0.42		✓
	HDL Cholesterol	NPHSII (98), EPIC (97), GRACE	0.75		✓
	sPLA2-IIA Levels	EPIC (97), UDACS (99), GRACE	4.66x10 ⁻⁷		✓
	sPLA2 Activity	EPIC (97), GRACE	0.87		✓
	CHD	NPHSII (98), EPIC (97), UDACS (99), GRACE	0.51		✓
rs11573203	LDL-cholesterol	NPHSII (97), UDACS (96), GRACE	0.03	UDACS Driven	✓
	HDL Cholesterol	NPHSII (97), UDACS (96)	0.13		✓
	sPLA2-IIA Levels	UDACS (96), GRACE	0.01		✓
	sPLA2 Activity	GRACE	0.93		✓
	CHD	NPHSII (97), UDACS (96), GRACE	0.79		✓

*Numbers in brackets represent the % of genotype calls made where results are available

**HWE = Hardy Weinberg Equilibrium. A tick represents that all studies genotyped for the SNP were within the HWE and not significantly different at 1 degree of freedom where data is available.

Table 5.2. A small initial meta-analysis showing the associations of non-lead *PLA2G5* tSNPs; rs11573248, rs11573203 and rs11573191 with relevant sPLA2 and CHD biomarkers. Yellow boxes indicates a significant association (P<0.05) (EPIC = EPIC-Norfolk, EPIC-NL= EPIC Netherlands).

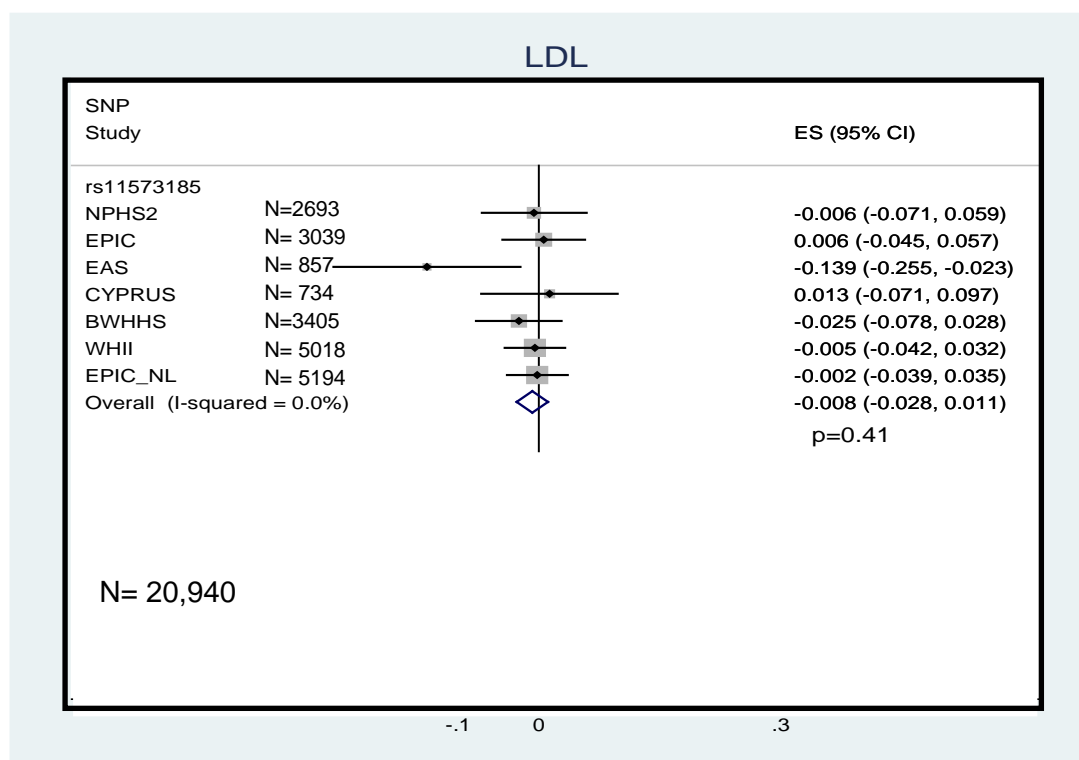


Figure 5.14. Adjusted Meta-analysis shows the association between LDL-cholesterol and rs11573185. ES = Effect Size. EPIC = EPIC-Norfolk, EPIC-NL = EPIC Netherlands studies. All studies defined previously in Table 2.6

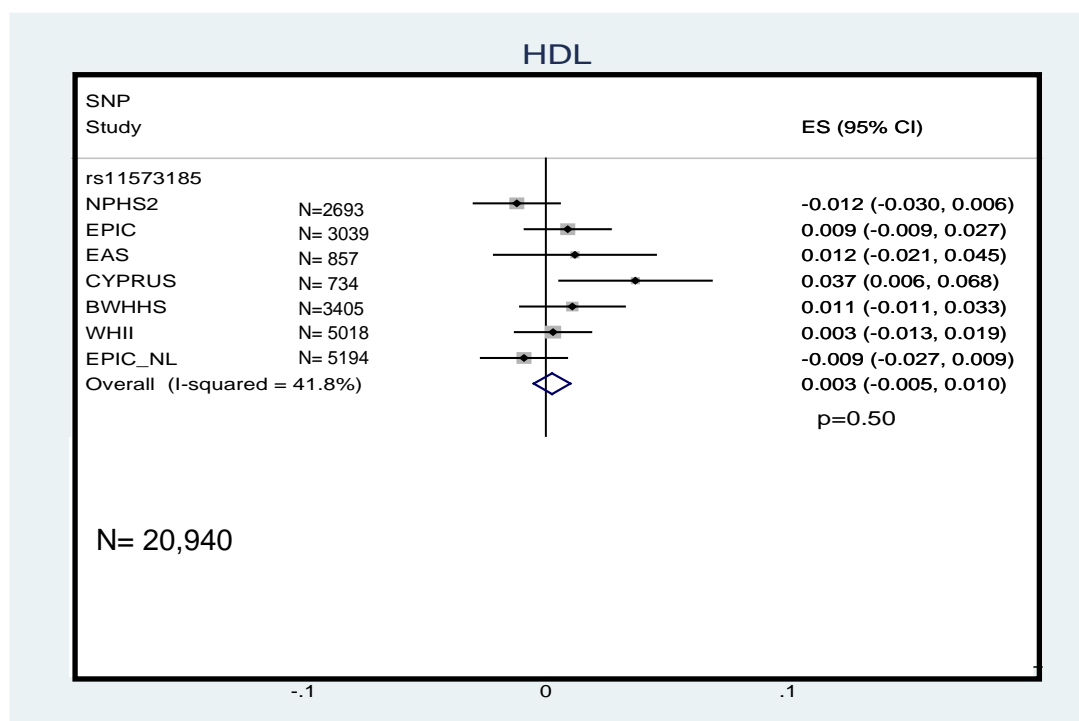


Figure 5.15. Adjusted Meta-analysis shows the association between HDL cholesterol and rs11573185. ES = Effect Size. EPIC = EPIC-Norfolk, EPIC-NL = EPIC Netherlands studies. All studies defined previously in Table 2.6

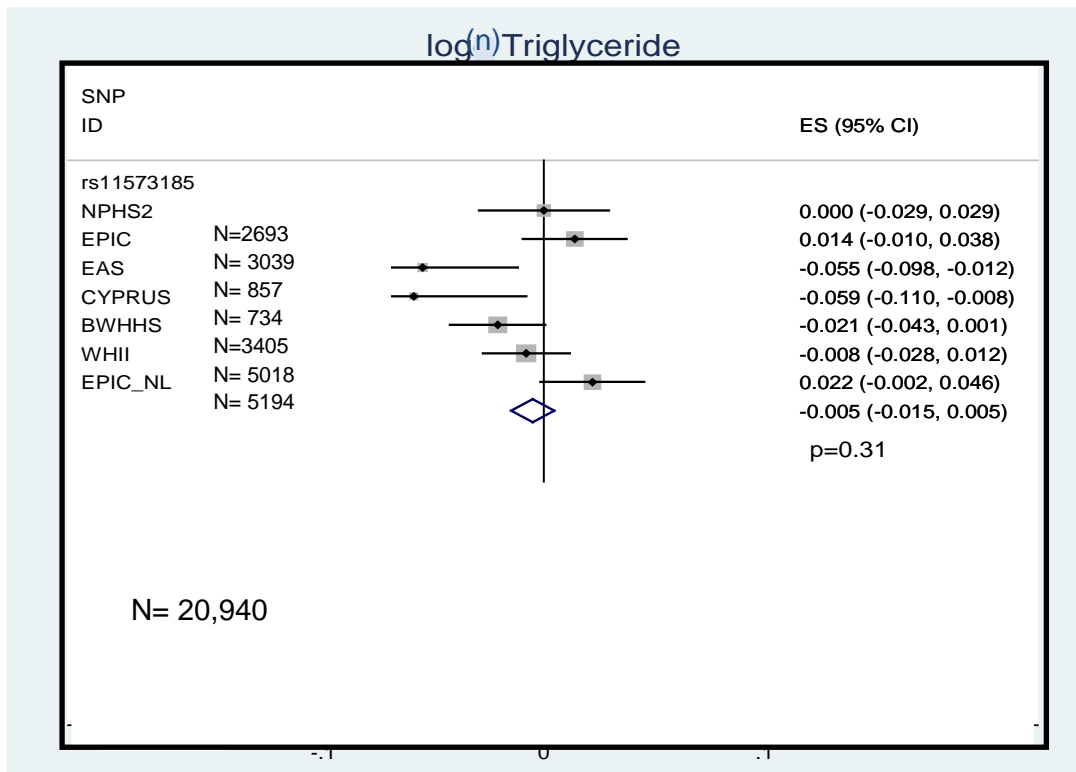


Figure 5.16. Adjusted Meta-analysis shows the association between (log) triglycerides and rs11573185. ES = Effect Size. EPIC = EPIC-Norfolk, EPIC-NL = EPIC Netherlands studies. All studies defined previously in Table 2.11

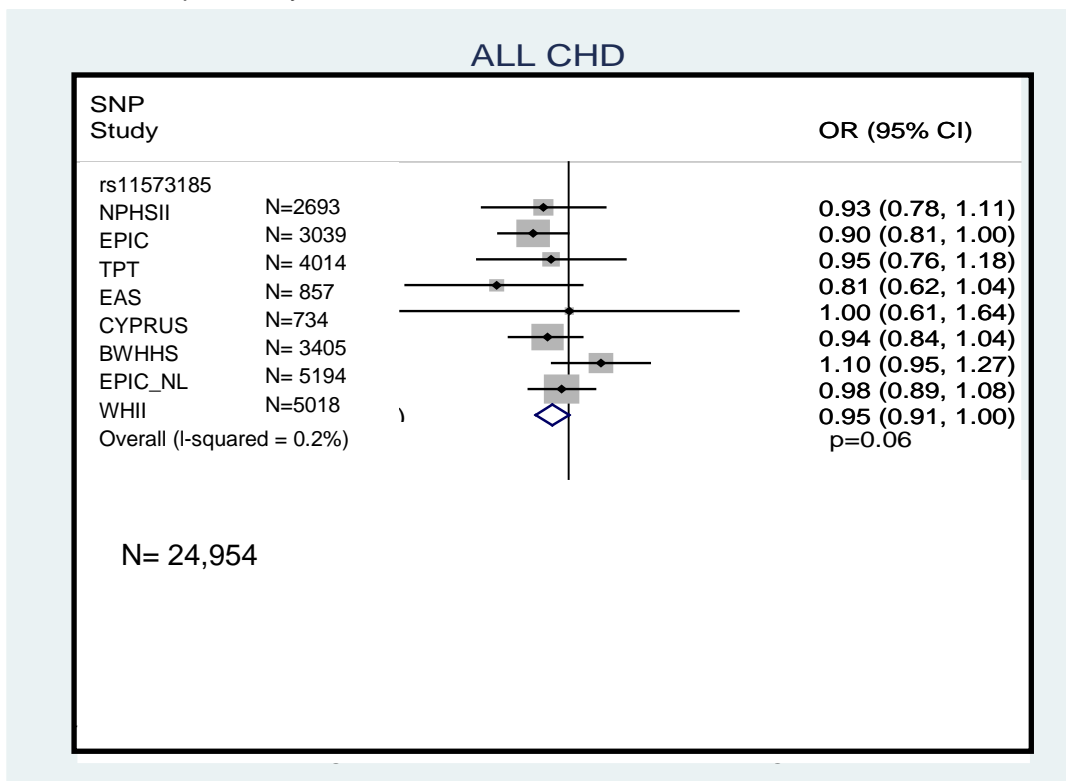


Figure 5.17. Adjusted Meta-analysis shows the association between CHD and rs11573185. OR = Odds Ratio. EPIC = EPIC-Norfolk, EPIC-NL = EPIC Netherlands studies. All studies defined previously in Table 2.6

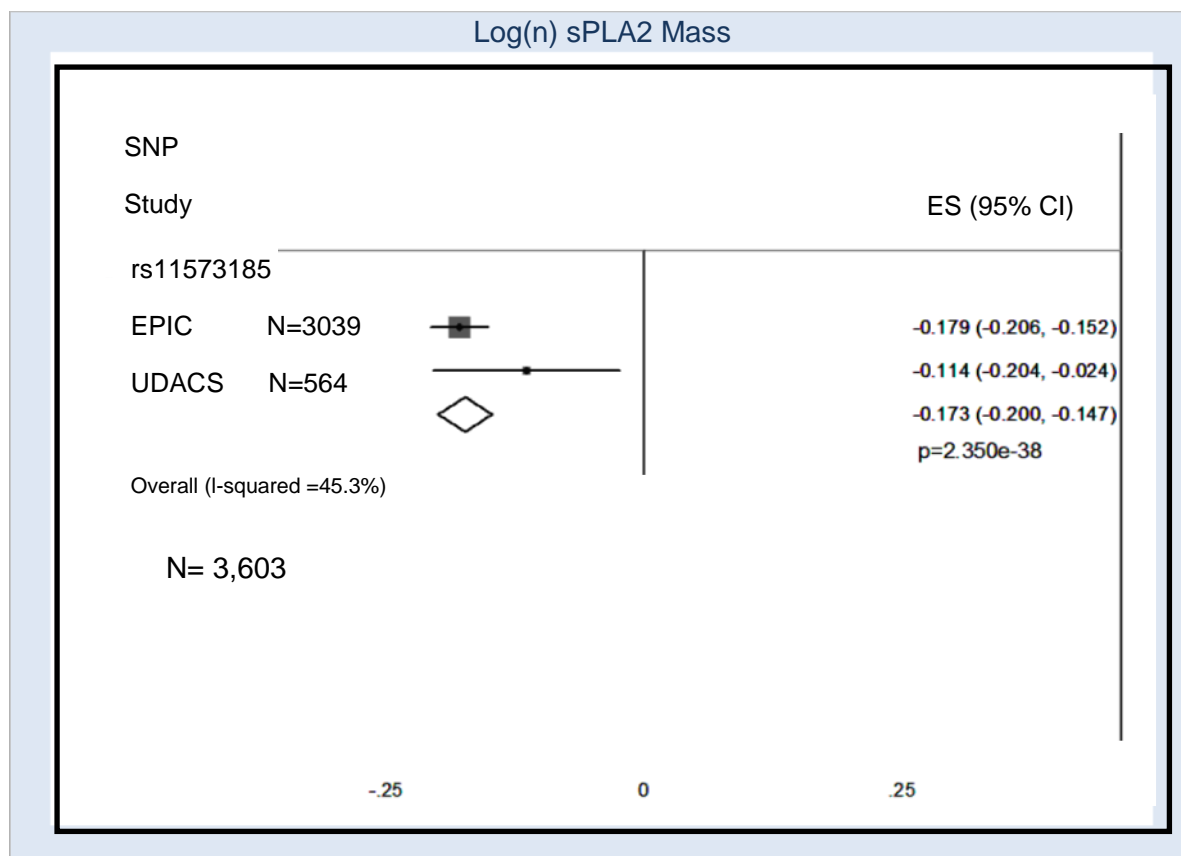


Figure 5.18. Adjusted Meta-analysis shows the association between sPLA2-IIA mass and rs11573185. ES = Effect Size. EPIC = EPIC-Norfolk. All studies defined previously in Table 2.11

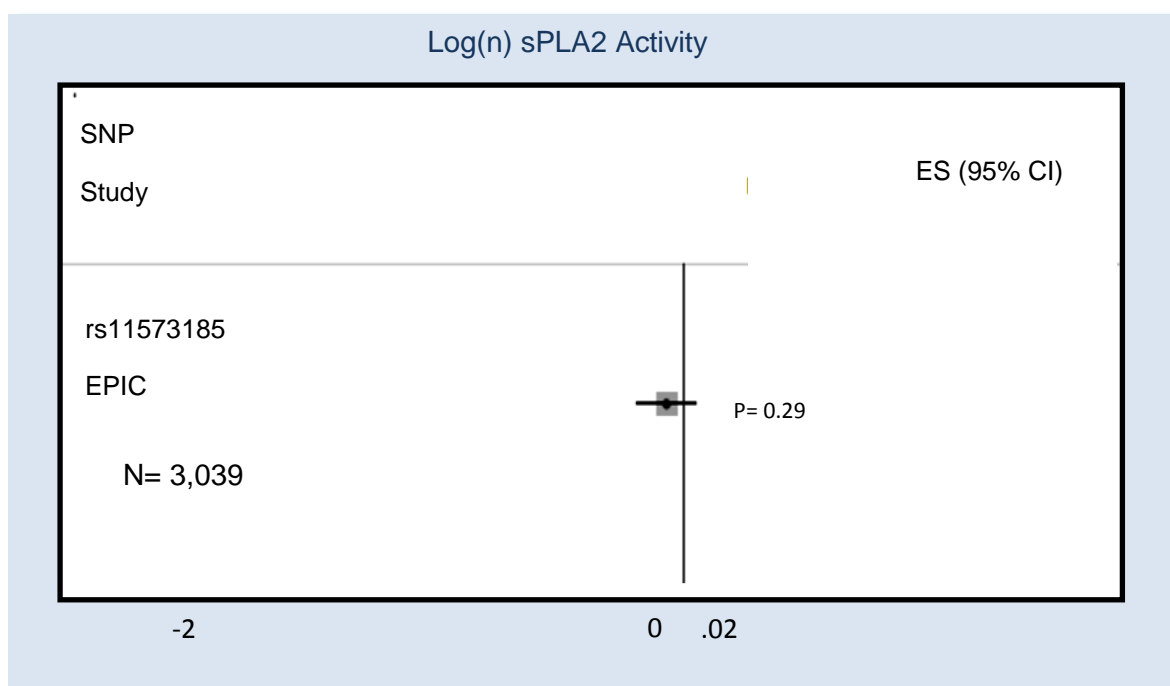


Figure 5.19. Adjusted analysis shows the association between sPLA2-activity and rs11573185. ES = Effect Size. EPIC = EPIC-Norfolk. This study was defined previously in Table 2.6

5.2.5. Combined Meta-Analysis Data for rs525380

The SNP rs525380, selected for further analysis due to its high per-SNP differential association with *PLA2G5* mRNA levels in the ASAP study was genotyped in EPIC-Norfolk. These data were added to the genotype data from the GWAS study, WTCCC. The results regarding the associations between rs525380, CHD, sPLA2-IIA mass and sPLA2 activity, are shown below (Figure 5.20-5.21).

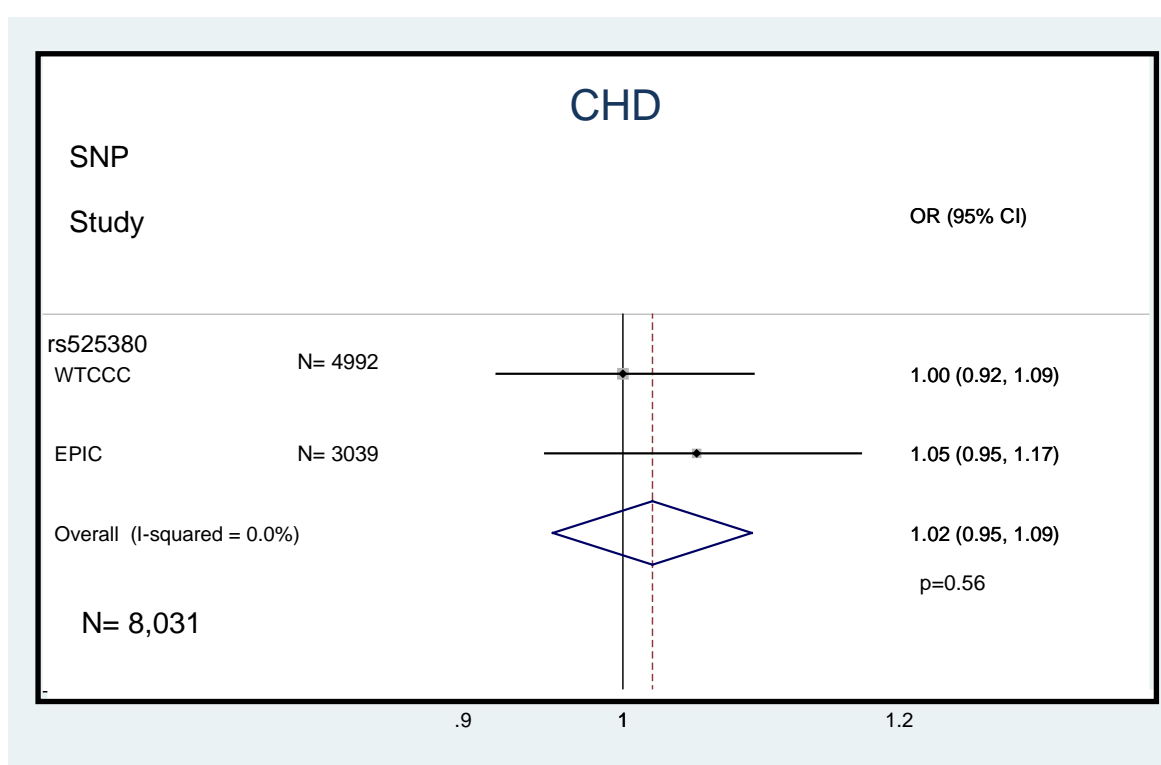


Figure 5.20. The per allele association of rs525380 with CHD. OR = Odds Ratio. EPIC = EPIC-Norfolk. All studies defined previously in Table 2.6.

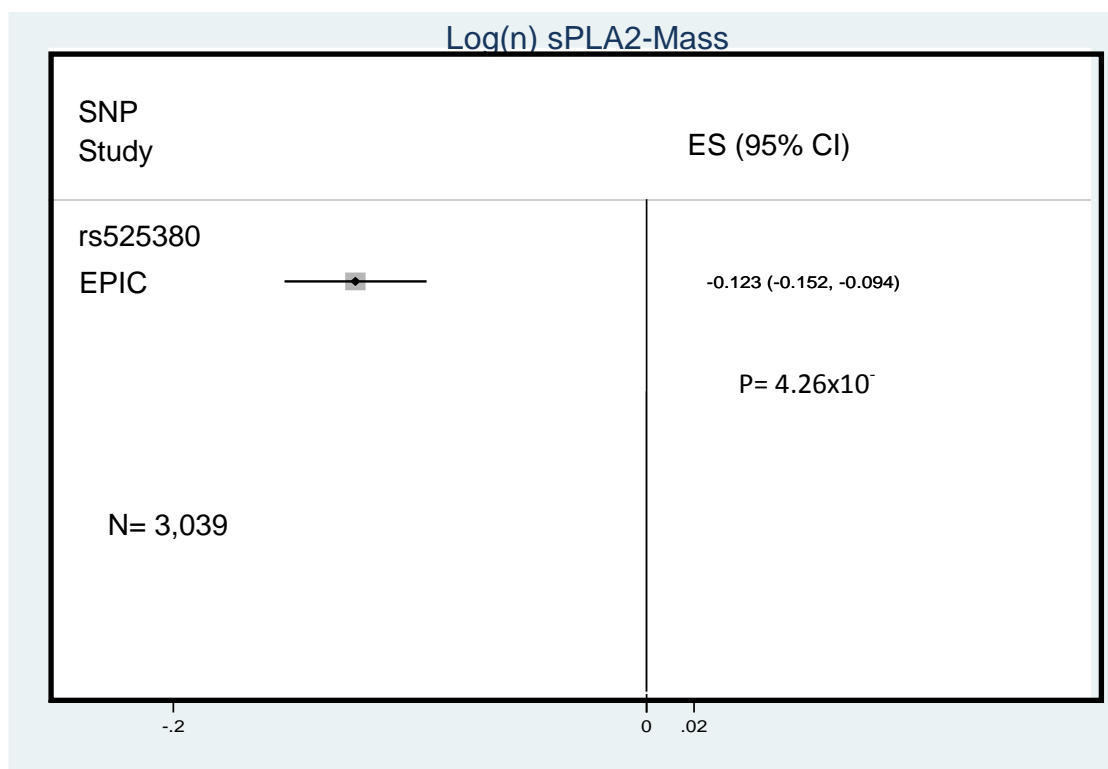


Figure 5.21. The association of rs525380 with sPLA2-IIA mass. ES = Effect Size. This study was previously defined in Table 2.6

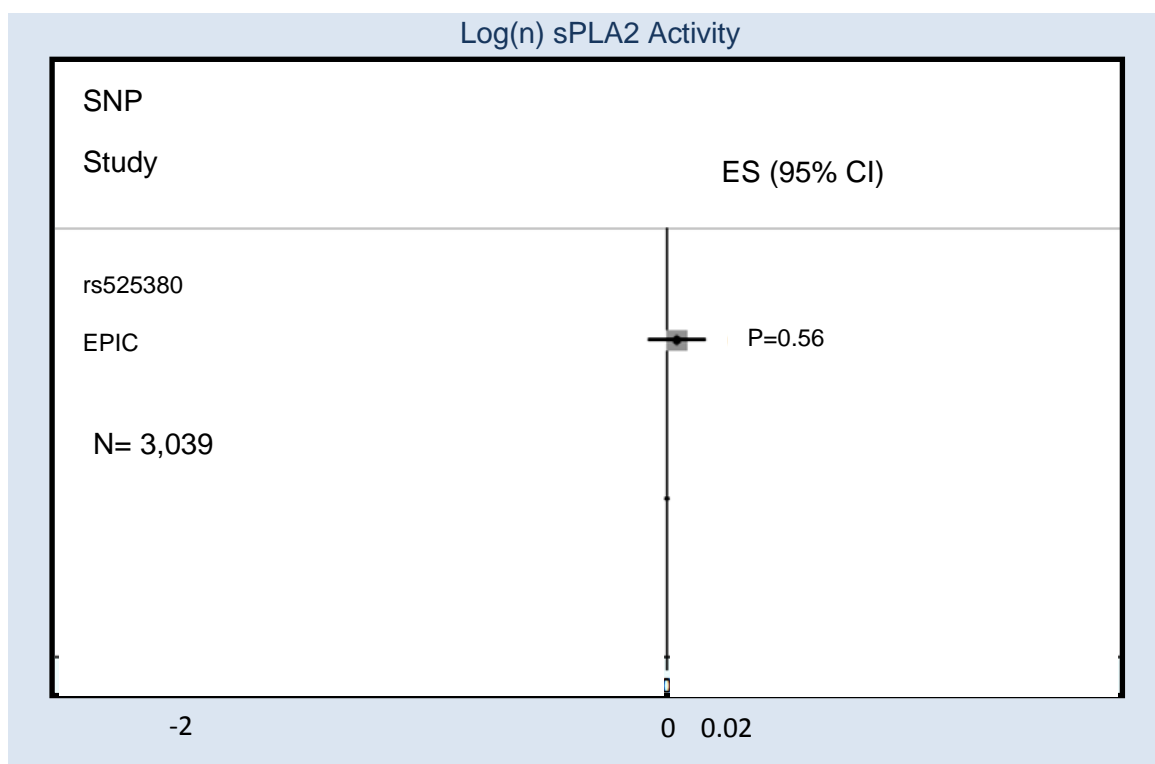


Figure 5.22. The association of rs525380 with sPLA2 activity. ES= Effect Size. This study was previously defined in Table 2.6

5.3. Discussion

The original tSNP selection and analysis in UDACS showed the associations of several tSNPs covering approximately 92% of the variance of the gene *PLA2G5* (Wootton, Arora et al. 2007). From this univariate analysis four SNPs; rs11573185 C>A, rs11573191 A>G, rs11573203 G indel and rs11573248 TA indel, were identified as having a potentially significant association with the CHD related biomarker, cholesterol and being significantly associated with sPLA2-IIA levels. These SNPs were genotyped in additional studies by me and others in my lab to try and establish whether these effects could be replicated.

Data from the meta-analysis indicated that none of the SNPs maintained a significant association with cholesterol and although rs11573203 showed a borderline significant association with LDL-cholesterol levels ($P=0.03$), these results were only inclusive of 3 and 2 relatively small cohorts respectively and in both instances the association was driven by the small UDACS cohort alone, suggesting that these may be anomalous results. Additionally, when taking into account the Bonferroni Effect it is unlikely that these results will remain significant.

Analysis of these tSNPs with sPLA2-IIA levels and sPLA2 activity showed that all SNPs were significantly associated with sPLA2-IIA levels, with a per rare allele increase seen with rs11573248 ($P=1.36 \times 10^{-40}$) and a per rare allele decrease seen with rs11573185 ($P=2.35 \times 10^{-38}$) being the most significant associations. However none of the SNPs were significantly associated with sPLA2 activity and none showed a significant association with CHD.

Regarding rs525380 A>C, the SNP selected from the ASAP study results as having the most significant differential effect on *PLA2G5* mRNA expression (in the aortic adventitia), genotyping analysis in the EPIC-Norfolk and WTCCC studies showed no significant associations with either sPLA2 activity or CHD, though there was a significant association between rs525380 and sPLA2-IIA levels ($P=7.74 \times 10^{-7}$). There is no reported significant LD between any of these genotyped SNPs.

The results from the meta-analysis suggest that there is no causal association between sPLA2-V and CHD or CHD related traits with any *PLA2G5* SNP analysed in this chapter. These results suggest sPLA2-V is likely to be an inflammatory marker of atherosclerosis, but not causally associated. The presence of sPLA2-V in plaque tissues (Jonsson-Rylander, Lundin et al. 2008) is therefore most likely to be due to reverse causation. However, the power of these results is relatively low due to limited studies and is being increased in order to confirm these findings.

Regarding the potential contribution of sPLA2-V to total sPLA2 activity, analysis of the results from the meta-analysis show no significant association between any of the genotyped *PLA2G5* SNPs and sPLA2-activity. Even in these relatively small cohorts, an effect on sPLA2 activity was seen with *PLA2G2A* SNPs (**Chapter 4**), therefore it seems unlikely that a stronger effect will be seen with any of the *PLA2G5* SNPs, even in a larger cohort. This suggests that the contribution of sPLA2-V to sPLA2 activity could be less than the contribution made by sPLA2-IIA in CHD relevant tissues, including the liver (where some sPLA2 isoenzymes are expressed to move into the circulation).

The ASAP study represents a significant portion of the gene *PLA2G5*. Both available SNPs of interest (rs11573185 and rs525380) were shown to be associated with differential *PLA2G5* mRNA expression in the aortic adventitia. However, rs11573185 was not significantly associated ($P=0.08$) whereas rs525380 was the single strongest association of a SNP with differential *PLA2G5* mRNA expression in this tissue ($P=5.50 \times 10^{-6}$), suggesting that rs525380 is strongly associated with differential sPLA2-V activity, but that rs11573185 may be less so. This supports the idea that rs525380 is strongly associated with differential sPLA2-V activity. However, since rs525380 was not shown to be significantly associated with total sPLA2 activity in the meta-analysis perhaps sPLA2-V does not contribute as strongly to overall sPLA2 activity measures compared to sPLA2-IIA in these CHD-related tissues in the circulation. Analysing selected SNPs of *PLA2G5* showed that the strongest overall differential expression was consistently in the aortic adventitia and liver, indicating that sPLA2-V is most highly expressed in these tissues, however when compared to the *PLA2G2A* SNPs, reported in **Chapter 3**, the highest expression levels of *PLA2G5* are approximately third as high as those of *PLA2G2A* ($-\log_{10}(P)$ values of 17.5 for *PLA2G2A* and 5.5 for *PLA2G5*). This suggests that the relative levels of sPLA2-IIA are considerably higher than sPLA2-V in CHD related tissues, as has been previously shown (Lambeau and Gelb 2008). Since data in this chapter shows sPLA2-V levels are comparatively lower and no association is seen with sPLA2 activity in these tissues and in the circulation, the 20 times higher potency associated with sPLA2-V compared to sPLA2-IIA *in vitro* and in animal models may not be significant in normal human physiological conditions for these tissues and in the circulation. These results also show that the associations of *PLA2G2A* and *PLA2G5* SNPs with mRNA expression are different depending on the tissue. *PLA2G2A* differential mRNA expression is seen most frequently and strongly in the liver, but for *PLA2G5* it is mostly in the aortic adventitia, suggesting that individual isoenzyme contributions to sPLA2 activity are potentially tissue dependent and based on expression patterns (Lambeau and Gelb 2008). Therefore, perhaps it is more likely that sPLA2 isoenzyme contributions to sPLA2 activity are flexible rather than fixed, depending on the

environment and type of tissue being studied. However, it should be noted that the power of the meta-analysis results is low due to the small number of studies included and is being increased in order to confirm the findings regarding sPLA2 activity. Also, as there are currently no specific ELISAs available to measure the levels of sPLA2-V or the specific contributions of each isoenzyme to total sPLA2 activity it is uncertain that this interpretation of the sPLA2-IIA and sPLA2-V contributions to activity is true. Results from the ASAP study analysing each SNP individually suggests there is an overlap between some of the *PLA2G5* associated SNPs and *PLA2G2A* mRNA expression in the liver, and significant meta-analysis associations have been shown with both rs11573185 and rs525380 (as well as the additional tSNPs rs11573191, rs11573203 and rs11573248) and sPLA2-IIA levels. These findings appear to add weight to the argument that sPLA2-IIA and sPLA2-V may be able to partially compensate for and independently regulate each other to a certain degree, despite their mechanistic differences (Sawada, Murakami et al. 1999). This may be unique to these enzymes, but could possibly be a feature of the whole *PLA2* gene-cluster on chromosome 1 and a potentially important continuation of this work would be to analyse the relationships of all the genes in this cluster and identify both the co-operative and independent activities between them. This overlap in regulation amongst these genes may also go some way to explaining why it is has been, thus far, so difficult to create an assay to measure the individual activity of each enzyme in the serum, since, in the case of sPLA2-IIA and sPLA2-V, the regulation and expression of these enzymes may be more tightly linked than first thought. It may be that the effects seen here are due to the functional effects of another SNP in high LD with rs11573185 or rs525380 within the 90kb, shared promoter region between *PLA2G2A* and *PLA2G5* (potentially the true functional SNP may not yet be genotyped, since I have been unable to identify any SNPs in high LD with rs11573185 in either the HapMap or 1000 Genomes databases). Other SNPs in this region that are not genotyped may not yet have been identified as specific for either gene. This shared promoter region and potential co-regulation between *PLA2G2A* and *PLA2G5* suggests that the tissue dependent expression patterns of each sPLA2 isoform may reflect the best suited enzyme for each individual tissue and yet the structural similarities between sPLA2-IIA and sPLA2-V mean that in each instance, should the isoform of choice be absent or degraded for some reason in a particular tissue, the second sPLA2 isoform could at least partially compensate for it. A definitive conclusion regarding these associations cannot be confirmed until the power of the meta-analysis is increased and, therefore, the association is tested more robustly. It is however also possible that the associations of these two *PLA2G5* SNPs with *PLA2G2A* mRNA expression could have a potentially confounding effect on the results in this chapter. For example, because sPLA2-V also alters sPLA2-IIA mRNA and data in Chapter 4 suggests that sPLA2-IIA is not causal for CHD, the null association between sPLA2-IIA with

CHD may dilute the effect of sPLA2-V on CHD towards null and result in negative confounding.

Although rs525380 appears to be a reasonable candidate gene to use in MR analysis, this must be confirmed by functional work, similar to what has been presented in **Chapter 3** to test whether it is functional or tags a truly functional SNP. We must also consider the problem of potential confounding by sPLA2-IIA levels and how this can be assessed and overcome.

A limitation of this study is that the tSNPs selected were chosen in 2007 and this selection method has now been updated by more advanced algorithms that are much more accurate at identifying tSNPs. Additionally, advances such as the completion of the 1000 genomes database have also improved our ability to identify SNPs of interest. Therefore it may be that there are other useful potential candidate SNPs for MR analysis besides those presented in this chapter.

5.4. Conclusion

Based on the results shown in this chapter, I cannot conclusively identify whether sPLA2-V contributes significantly to sPLA2-activity compared to sPLA2-IIA in CHD related tissues using this genetic approach. Regarding the small meta-analysis I would suggest that rs11573185 (associated with sPLA2-IIA levels) and rs525380 (associated with *PLA2G5* mRNA expression) as well as rs11573203 (borderline association with LDL and triglycerides) and rs11573248 (associated with sPLA2-IIA levels) should be genotyped in additional studies, as increasing the power of the meta analysis will allow us to be confident in making a true conclusion. It may also be prudent to further explore the relationships between all of the genes identified as part of the *PLA2* cluster on chromosome 1 to best interpret how they may partially share regulatory mechanisms.

CHAPTER 6

The Potential Contribution of sPLA2-X to sPLA2 Activity

6.1. Introduction

6.1.1. sPLA2-X Background

Previous investigation of all the mammalian sPLA2s has shown that sPLA2-X exhibits the strongest PC-hydrolysing activity (Karabina, Brocheriou et al. 2006). This is potentially important when evaluating the contribution of this enzyme to both overall sPLA2 activity and CHD risk. Chapter 1.5 gives an overview of sPLA2-X and explores how sPLA2-X differs from sPLA2-IIA and sPLA2-V. The gene for sPLA2-X is *PLA2G10* which exists at chromosome 16p13.1-p12, not as part of the *PLA2* gene cluster on chromosome 1.

In this Chapter I used genotyping data from a small meta-analysis and eQTL data from the ASAP study to explore the potential contribution of sPLA2-X to total sPLA2 activity and to test the possibility that sPLA2-X acts as a causal biomarker for CHD. Identification of potential functional SNPs allowed me to assess whether any of these variations have significant associations to sPLA2-IIA mass, sPLA2 activity or CHD, by genotyping across several studies previously identified (2.5). The eQTL data from the ASAP study then allowed me to analyse how *PLA2G10* mRNA levels compare to both *PLA2G2A* and *PLA2G5* mRNA expression in CHD related tissues such as liver, aorta adventitia, aorta media and heart. This suggests what the potential contribution of sPLA2-X might be to total sPLA2 activity.

6.1.2. Identifying Tagging-SNPs of Interest

As described in 1.7.2, 10 tSNPs were identified by the STRAM algorithm and three of these SNPs with an MAF >5% were selected for further analysis by genotyping in both EPIC-Norfolk and GRACE-France: rs72546340, rs72546339 and rs4003232. A further tSNP, the C-SNP rs4003228 (R38C) was identified in the published literature (Gora, Perret et al. 2009), and was also included in my analysis despite a low reported MAF of ~3%. There is ambiguous mapping of rs4003228 in the bioinformatics databases I used to analyse these SNPs, due to a lack of genotyping information in relevant CEU populations. The SNP has, however, been mapped in published literature (Gora, Perret et al. 2009). The *PLA2G10* gene was screened by this group using the genomic sequences retrieved from public depositories and performed using genomic DNA from 62 unrelated MI patients selected from the ECTIM

study (<http://www.genecanvas.org>). Researchers used polymerase chain reaction (PCR) and direct sequencing, to examine 1 kb of the promoter, all of the exonic sequences and their corresponding flanking intronic sequences, as well as 1 kb of the 3'UTR downstream of the codon stop. The sequences were aligned and analysed with the SeqScape software (Applied Biosystems) to determine the potential polymorphisms. The genotyping of the polymorphisms were carried out using a modified RFLP method or TaqMan 5' nuclease assays. Details of the polymorphisms and genotyping conditions is described in detail on the group's website web site (<http://www.genecanvas.org>, genes: *PLA2G10*).

6.2. Results

6.2.1. Bioinformatics Analysis of *PLA2G10*

Data from the Ensembl Genome Browser confirms that *PLA2G10* is located at 16:14,766,405-14,788,523. The gene is reverse transcribed and encodes the enzyme sPLA2-X. According to the Ensembl Genome Browser and NCBI databases there are 2 known transcripts of the gene and one of these codes for the normal protein. The second is subject to nonsense mediated decay, where a premature stop codon is recognized and undergoes a decay process preventing the sequence from translating a normal protein. Evidence for this is presented by the Ensembl Genebuild and Havana Projects including protein, EST, and cDNA sequence data in support of both the full length and nonsense mediated decay transcripts for *PLA2G210*.

The protein coding transcript is made up of 4 exons (Figure 6.1a). This transcript codes for a 165 amino acid protein. The nonsense-mediated decay truncated transcript (Figure 6.1b) is 587bp, producing a 130 amino acid protein.

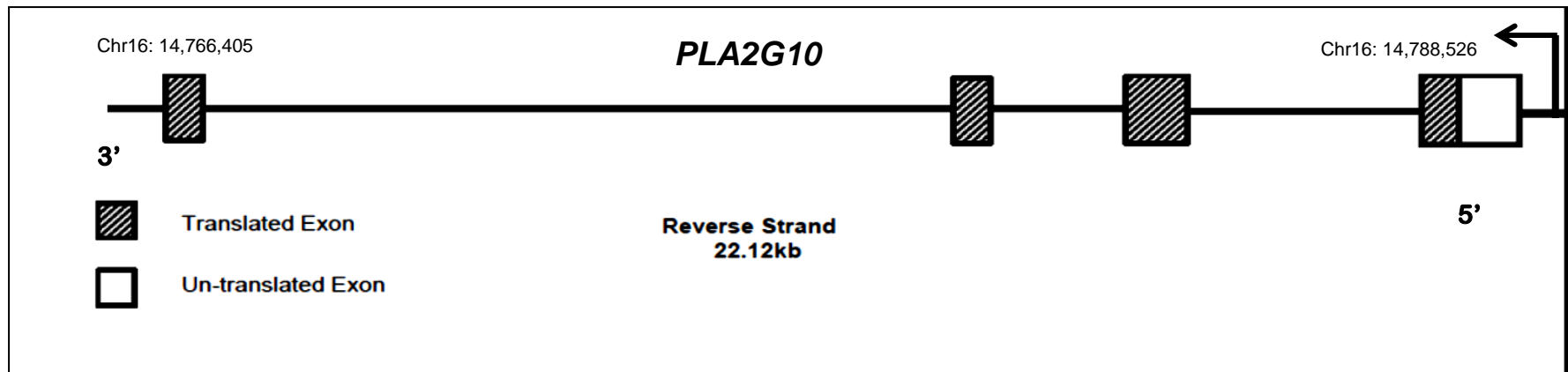


Figure 6.1a. *PLA2G10* coding transcript (Ensembl Transcript ID: ENST00000438167) for normal protein. This schematic is taken from the Ensembl Genome Browser. The boxes indicate exons and the solid line represents introns.

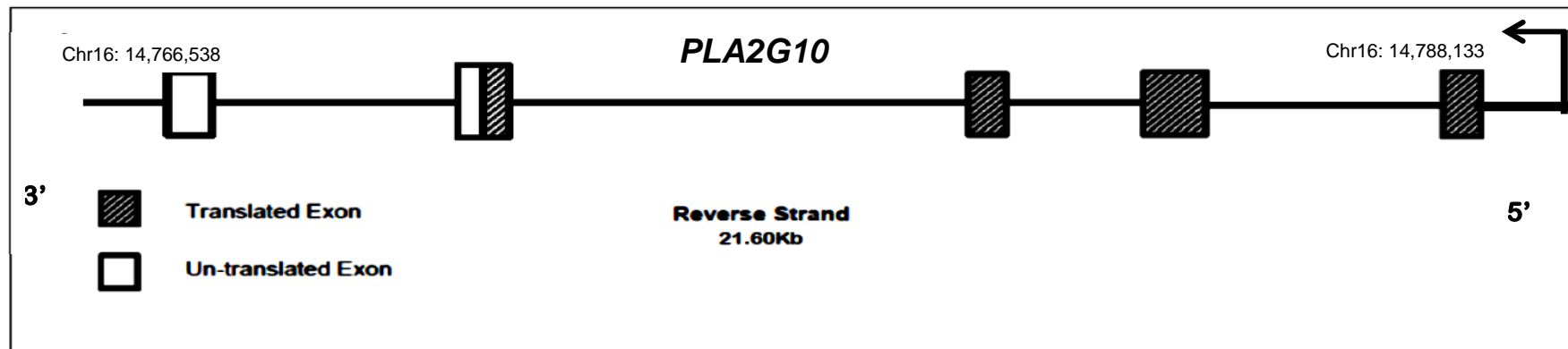


Figure 6.1b. *PLA2G10* transcript subject to nonsense-mediated decay (Ensembl Transcript ID: ENST00000567462) for truncated protein. This schematic is taken from the Ensembl Genome Browser. The boxes indicate exons and the solid line represents introns.

6.2.2. Results from EPIC-Norfolk and GRACE-France

Tables 6.1, 6.2 and 6.3 show the associations between the four selected SNPs and sPLA2-IIA mass in EPIC-Norfolk and GRACE-France, sPLA2 activity in EPIC-Norfolk and GRACE-France and CHD in EPIC-Norfolk respectively. Only rs4003228 shows a significant association with any trait, exhibiting a significant association with sPLA2 activity in the GRACE-France cohort. All genotyping completed in EPIC-Norfolk and GRACE-France complied with the Hardy Weinberg Equilibrium (not significantly different at 1 degree of freedom).

<i>PLA2G10</i>		EPIC			GRACE		
SNP	Genotype	N	sPLA2 activity geometric mean (approximate SD)	MAF	N	sPLA2 activity geometric mean (approximate SD)	MAF
rs72546340 (G>C)	GG	2955	4.47 (1.12)	0.03	297	3.10 (1.13)	0.05
	GC	171	4.56 (1.08)		16	3.48 (1.20)	
	CC	9	4.18 (0.69)				
	P value (trend) ¹		0.53			0.23	
	P value ²		0.52			0.25	
rs72546339 (C>A)	CC	2901	4.47 (1.13)	0.04	294	3.14 (1.16)	0.03
	CA	279	4.45 (1.04)		23	2.97 (1.03)	
	AA	13	4.37 (0.99)				
	P value (trend) ¹		0.74			0.51	
	P value ²		0.79			0.90	
rs4003228 (R38C)	CC	2151	4.47 (1.12)	0.02	301	3.17 (1.12)	0.02
	CT	90	CT+TT:4.42 (0.97)		15	2.47 (1.30)	
	TT	1			0	--	
	P value ¹		0.66			0.01	
	P value ²		0.59			0.005	
rs4003232 (C indel)	TT	1460	4.47 (1.06)	0.21	141	3.17 (1.21)	0.27
	TC	760	4.49 (1.19)		100	3.10 (1.05)	
	CC	110	4.54 (1.22)		22	3.02 (1.21)	
	P value (trend) ¹		0.52			0.52	
	P value ²		0.52			0.60	

All results are represented as geometric mean (approximate SD) 11: homozygotes for the common allele; 12: heterozygotes; 22: homozygotes for the rare allele. ¹ unadjusted P values for trend. ² age and gender adjusted P value for trend.

Table 6.1. Associations between rs4003228, rs4003232, rs72546339 and rs72546340 variants and the trait, sPLA2 activity in the case:control nested cohort EPIC-Norfolk and the ACS cohort GRACE-France.

PLA2G10			EPIC		GRACE		
SNP	Genotype	N	sPLA2-IIA mass geometric mean (approximate SD)		N	sPLA2-IIA mass geometric mean (approximate SD)	
rs72546340	11	2955	8.60 [5.96-13.12]		297	2.97 [1.62-6.52]	
	12	171	8.30 [6.30-14.6]		16	2.50 [1.74-13.76]	
	22	9	7.00 [4.40-8.09]				
	P value (trend) ¹		0.65			0.68	
	P value ²		0.65			0.61	
rs72546339	11	2901	8.60 [6.00-13.390]		294	2.79 [1.59-7.08]	
	12	279	8.50 [5.950-12.60]		23	4.27 [1.94-5.76]	
	22	13	10.20 [7.80-12.70]				
	P value (trend) ¹		0.91			0.36	
	P value ²		0.60			0.19	
rs4003228	11	2151	8.40 [5.90-12.80]		301	2.97 [1.65-6.78]	
	12	90	8.00 [5.30-12.80]		15	2.95 [1.54-6.21]	
	22						
	P value (trend) ¹		0.59			0.91	
	P value ²		0.46			0.97	
rs4003232	11	1460	8.60 [6.00-13.10]		141	3.48 [1.73-6.99]	
	12	760	8.20 [5.70-12.40]		100	2.63 [1.60-7.69]	
	22	110	7.90 [5.40-11.60]		22	2.40 [1.35-4.97]	
	P value (trend) ¹		0.11			0.21	
	P value ²		0.12			0.45	

All results are represented as median (IQR). 11: homozygotes for the common allele; 12: heterozygotes; 22: homozygotes for the rare allele. ¹ unadjusted P values. ² age, gender and CHD adjusted P value.

Table 6.2. Associations between rs4003228, rs4003232, rs72546339 and rs72546340 variations and the trait, sPLA2-IIA Mass in the nested case:control cohort EPIC-Norfolk and the ACS cohort GRACE-France.

PLA2G10 SNP	Genotype	Freq in controls	Freq in cases	OR (95% CI)*
rs72546340 (303G>C)	GG	1977 (94.4)	978 (94.0)	1.00
	GC+CC	117 (5.6)	63 (6.1)	1.09 (0.79-1.49)
	P value			0.60
rs72546339 (264C>A)	CC	1942 (91.4)	959 (89.8)	1.00
	CA	175 (8.2)	104 (9.7)	1.20 (0.93-1.55)
	AA	8 (0.4)	5 (0.5)	1.27 (0.41-3.88)
	P value			0.15
rs4003228 (R38C)	CC	1201 (96.00)	950 (96.00)	1.00
	CT+TT	50 (4.00)	40 (4.00)	1.01 (0.66-1.54)
	P value			0.98
rs4003232 (T-123/in1C)	TT	813 (62.40)	647 (62.90)	1.00
	TC	422 (32.40)	338 (32.90)	1.01 (0.85-1.20)
	CC	67 (5.20)	43 (4.20)	0.79 (0.53-1.17)
	P value			0.51

*age and gender adjusted odds ratio

Table 6.3. Associations between rs4003228, rs4003232, rs72546339 and rs72546340 and CHD in the nested case:control cohort EPIC-Norfolk.

6.2.3. Bioinformatics Analysis of rs4003228 (R38C)

The results above in **6.2.2** show only rs4003228 to have any associations with the traits of interest. Therefore rs4003228 was taken forward as the lead SNP to be examined in more depth. When analysing rs4003228 G>A (R38C), there is a certain amount of ambiguity since, like some other *PLA2G10* SNPs, it has been mapped to four different positions according to the Ensembl Genome Browser, the UCSC Genome Browser, and the 1000 Genomes database. These are; Chr16:14,771,169, chr16:15,452,850, chr16:16,492,437 and chr16:18,407,169. None of these databases definitively explain why there are 4 known SNP locations, but there is some suggestion that it could be due to multiple submissions of the SNP location from different sources that have analysed different transcripts of the gene or that there are additional transcripts of the genes besides those defined by supporting evidence that we already recognise (Figure 6.1a and 6.1b) which would cause the SNP to be mapped differently relative to the length of each transcript. The issue of multiple locations for this SNP is complicated due to a lack of genotyping in the populations of interest, CEU HapMap release #28 and 1000 Genomes pilot 1 or phase 1 data and the generally limited genotype data that is available (only 4 individuals across all databases). However, previous data have mapped rs4003228 R38C as described in **6.1.2**. Results showed the SNP to be mapped to exon 2 of *PLA2G10*, at position Chr16: 14,691,989. The minor allele frequency was low at 3% and predicted by the analysis of the *SIPLAC* (n= 312) and *AtheroGene* (n=1,303) studies this is supported by data from my completed EPIC-Norfolk and GRACE-France genotyping, which showed the frequency of the SNP to be ~4%. However, this is still a very small combined Caucasian population (Gora, Perret et al. 2009).

6.2.4. Meta-Analysis for rs4003228 (R38C)

In order to assess the value of rs4003228 as a functional SNP that is differentially associated with sPLA2 activity and as a useful tool to analyse the relationship of sPLA2-X with CHD risk, Dr Montse Guardiola and I genotyped the variant in two additional studies, making up a small meta-analysis of 4 studies, 3 general population and 1 ACS cohort. The *AtheroGene* and *SIPLAC* results were then also combined for analysis of any associations between rs4003228 and CHD.

After adjustment for multiple variables including age, sex and smoking, the effect size (ES) for rs4003228 with (log) sPLA2-IIA mass was -0.019 (95% CI: -0.134, 0.095) (P=0.74), i.e. showing no significant association with mass (Figure 6.2). Likewise with (log) sPLA2 activity the effect size was shown to be non-significant at -0.022 (95% CI: -0.074, 0.030) (P=0.42) (Figure 6.3). These data suggest that despite the functional analysis of previous studies (Gora, Perret et al. 2009) showing that sPLA2-X activity is affected by this SNP *in*

vitro, the SNP does not manifest itself at a significant level to be relevant in this small scale meta-analysis. This is potentially due to the low MAF of this SNP (3%) in the general population.

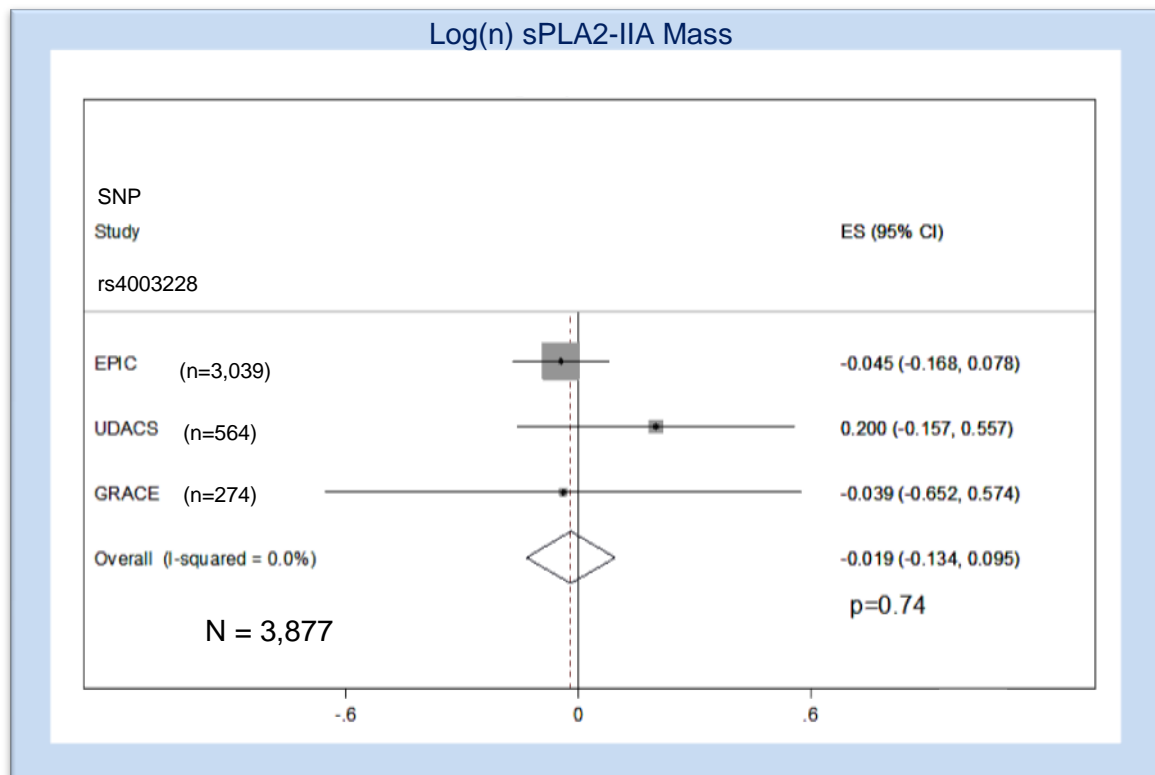


Figure 6.2. The association of *PLA2G10* SNP rs4003228 with sPLA2-IIA Mass. ES= Effect size. Studies have been previously defined (Table 2.6)

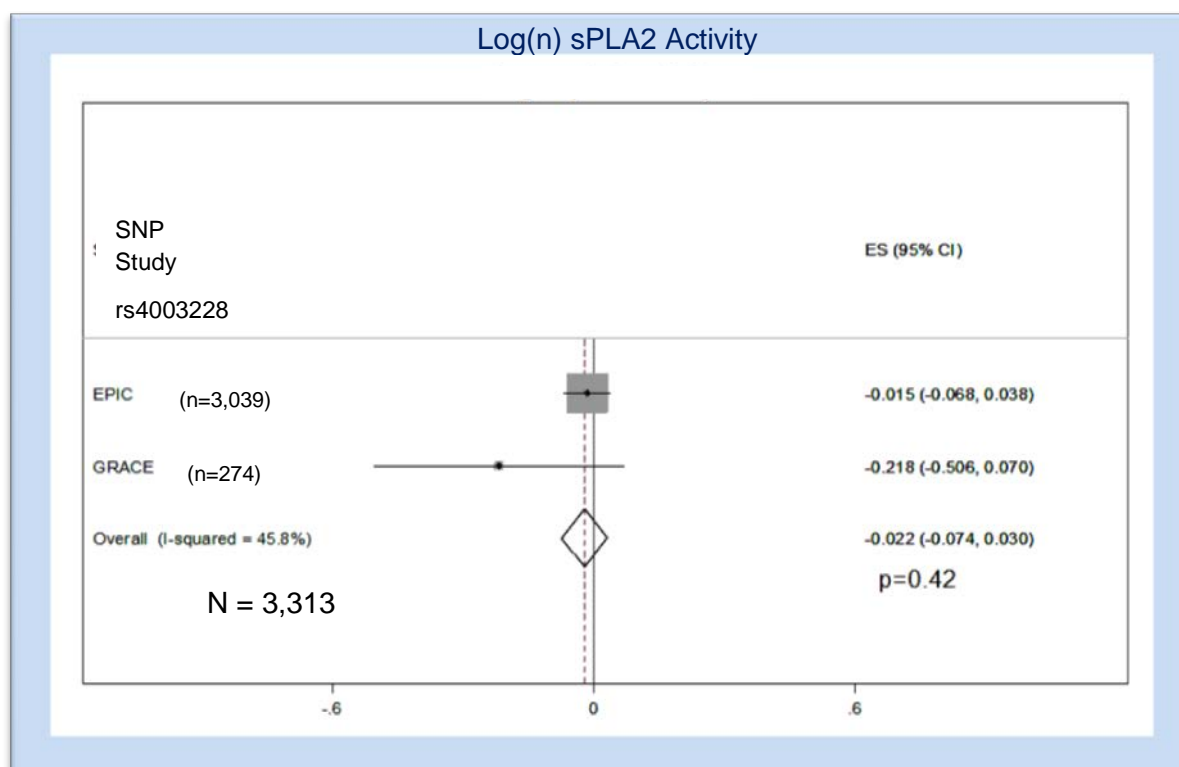


Figure 6.3. The association of *PLA2G10* SNP rs4003228 with sPLA2 Activity. ES= Effect size. Studies have been previously defined (Table 2.6)

Regarding disease (binary) associations, rs4003228 showed no association with CHD. After adjustment for age and gender, the odds ratio (OR) for the association of rs4003228 with CHD was 1.20 (95% CI: 0.89, 1.62) (P=0.24) (Figure 6.4). GRACE-France was not included in CHD measures since all events in this study are recurrent and may skew the data.

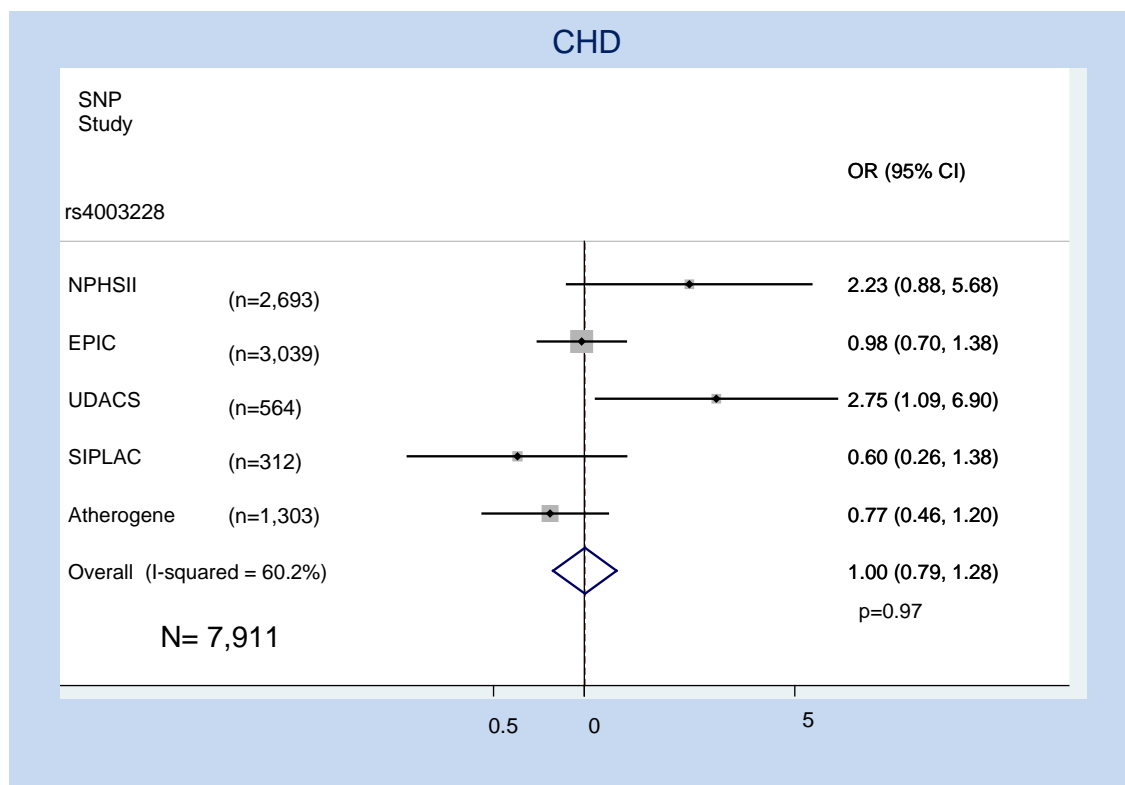


Figure 6.4. The association between *PLA2G10* SNP rs4003228 and CHD. OR = Odds ratio. Studies previously defined (Table 2.6).

6.2.5. *PLA2G10* mRNA Expression in the ASAP Study

The analysis I requested and analysed from Dr Folkersen regarding the expression of *PLA2G10* mRNA in various tissues showed that values were low in comparison to *PLA2G2A* (Figure 6.5 compared to Figure 4.4A) in these tissues. None of our selected SNPs of interest were present on the ASAP arrays and a limitation of these results is that only a small number of *PLA2G10* specific SNPs (13) were available on the array. However, this data does suggest that *PLA2G10* mRNA does not appear to be differentially expressed by any other *PLA2* related SNPs, suggesting that there is no potential for co-regulation between *PLA2G10* and other *PLA2* genes as we saw for *PLA2G2A* and *PLA2G5* in **Chapter 5**.

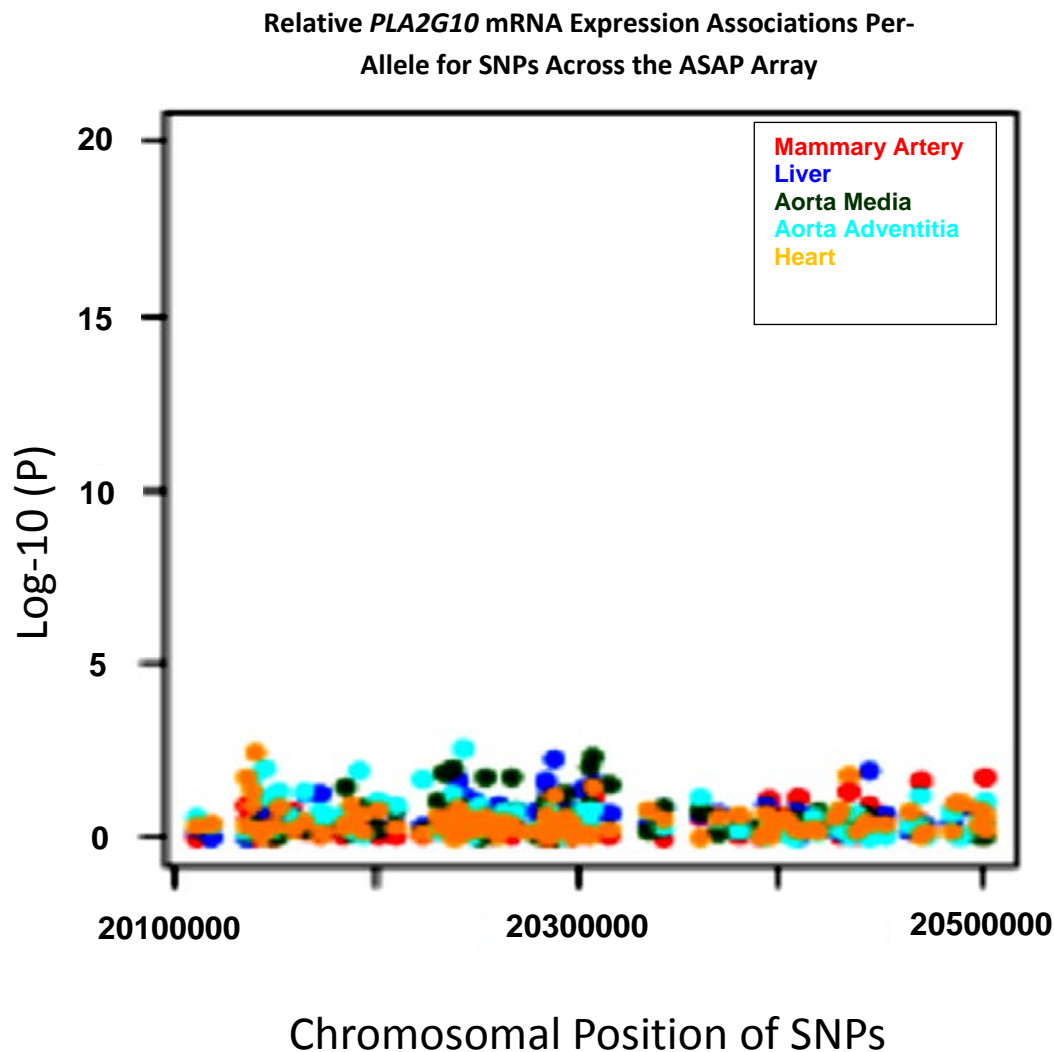


Figure 6.5. Relative expression levels of *PLA2G10* mRNA in various tissues of the ASAP study (Indicated by colour key). No SNPs appear to be significantly associated with differential *PLA2G10* mRNA.

6.3. Discussion

In order to begin to examine the contribution of sPLA2-X to the composite measure of sPLA2 activity and its potential contribution to CHD risk I took a two-pronged approach, first assessing a SNP previously identified as potentially affecting sPLA2-X in a small meta-analysis for any associations between sPLA2 activity and, potentially, CHD and then using data I had requested from the ASAP study to examine the mRNA expression of *PLA2G10* in several tissues.

The second transcript of *PLA2G10* identified by the Ensembl Genome Browser, indicates that nonsense mediated decay prevents the sequence from fully translating. The

variation thought to cause this premature stop codon is rs76137801C>G. The prediction algorithm websites SIFT and Polyphen predict that the change causes a residue change from R>G and is 'deleterious' and 'probably damaging' respectively. This SNP was not in the list of polymorphisms identified as being of interest relative to sPLA2-X (Gora, Perret et al. 2009) (STRAM) that I used to select my sPLA2-X analysis instrument and an interesting follow up to this chapter could be to genotype this SNP and observe its effect on the expression of sPLA2-X in different individuals and whether there is an association with this SNP that could perhaps explain low levels of *PLA2G10* mRNA expression detected in the ASAP study.

The results from the ASAP eQTL analysis show the expression of *PLA2G10* across all the examined tissues is very low and no single *PLA2G10* SNP on the array had a significant association with differential mRNA expression. This fits with data from previous studies showing sPLA2-X to be less frequently expressed than sPLA2-IIA and V in tissues such as the heart and liver (Kudo and Murakami 2002). The increased potency of sPLA2-X suggests that this enzyme should perhaps contribute strongly to total sPLA2 activity. However, the results from the ASAP study suggest that the protein levels expressed in these tissues would be so low that they are almost undetectable and despite the around 20 times higher potency of sPLA2-X compared to sPLA2-IIA, these low levels would very likely prevent any significant contribution to overall sPLA2 activity. This data cannot, however, give us any indication of the *PLA2G10* mRNA expression levels in tissues not measured by the ASAP study. Previous studies suggest that sPLA2-X is most strongly expressed in tissues other than those measured in the ASAP study, for example in the airways (Hallstrand, Chi et al. 2007; Henderson, Chi et al. 2007) and testes (Escoffier, Jemel et al. 2010). I think that this data supports the theory that individual isoenzyme contributions to sPLA2 activity are likely to be tissue dependent depending on the specific tissue requirements and the immediate environment.

Regarding the meta-analysis results, none of the previously identified SNPs showed significant associations with sPLA2 activity measures, suggesting that either these SNPs are not significantly associated with sPLA2-X or that sPLA2-X is not contributing strongly to total sPLA2 activity. A limitation of our study is that the STRAM method used to select *PLA2G10* SNPs of interest may now be outdated by technology such as genome wide association studies and therefore there are potentially better candidate SNPs available. However, we also took into account data from other groups previous studies (Gora, Perret et al. 2009) and their SNP selection and were still unable to show a significant association between sPLA2 activity and a sPLA2-X related SNP.

It is possible that sPLA2-X is expressed due to induction as part of the inflammatory response. This would help to explain why there is no association seen between sPLA2-X related SNPs and CHD, since *PLA2G10* appears not to be expressed at high levels in the normal artery, but is identified as being present in the atherosclerotic plaque (Hanasaki, Yamada et al. 2002; Murakami and Kudo 2003), suggesting that it is not causal for CHD but is initiated as part of the inflammatory response associated with atherosclerosis (Ross 1999).

This lead SNP, rs4003228 R38C was identified as potentially having an effect on sPLA2-X activity *in vitro* and therefore CHD in the study by Gora et al (Gora, Perret et al. 2009). The functional analysis of this SNP suggests that the synonymous arginine to cysteine change at codon 38 has a very significant impact on sPLA2-X activity *in vitro* with the extra cysteine residue causing ~95% decrease in sPLA2-X protein in COS cell supernatant and ~85% decrease in COS cell lysate. However, this SNP did not show any association to CHD risk in either the *AtheroGene* or *SIPLAC* studies. In order to try and establish whether this SNP has an effect on sPLA2 activity in human serum or on CHD risk we genotyped it in a further 4 studies and, as discussed above, this small combined meta-analysis resulted in no significant associations with either sPLA2-IIA mass or sPLA2 activity as well as showing no significant association with CHD. The meta-analysis is not highly powered and the low minor allele frequency of the SNP (3-4%) suggests that a larger cohort would need to be established before any firm conclusions can be drawn about this SNP regarding sPLA2-X activity and therefore whether it is a useful genetic instrument for further MR studies to investigate the role of sPLA2-X in CHD. However, from these results, we can see that although the SNP is shown to have a significant effect *in vitro* (Gora, Perret et al. 2009), it does not appear to affect sPLA2 activity *in vivo*, suggesting that perhaps the SNP is not having an impact in the general population due to its low frequency, or that the SNP is significantly associated to sPLA2-X activity, but that sPLA2-X is generally not well expressed and does not contribute significantly to sPLA2 activity in organs such as the heart and liver, despite its potency. This would mean that regardless of a significant association with group X sPLA2 activity, the association between the SNP and total sPLA2 activity would be low. This would also fit with the evidence that sPLA2-X is expressed at low levels in these tissues compared to other sPLA2 isoforms (Kudo and Murakami 2002). The low power of the study is a limitation and a larger study would need to be conducted to increase the power of these findings. Unfortunately, due to time constraints, I was unable to follow up this initial meta-analysis, but strongly suggest additional work to further validate these results.

As to other potential variants that could affect sPLA2-X activity, these have proved difficult to identify due to ambiguous mapping to chromosomal positions and a subsequent

lack of known genotyping or functional data. Again future work should consider the potential of other *PLA2G10* SNPs as better instruments (with higher MAFs) to study the effects of sPLA2-X on CHD risk and its contribution to overall sPLA2 activity, perhaps identified through platforms such as GWAS.

PLA2G10 is not part of the *PLA2* cluster on chromosome 1 and thus does not contribute to any potential co-regulation between the cluster genes, despite structural similarities with these enzymes (Murakami and Kudo 2003). Transgenic mice over-expressing *pla2g10* die neonatally due to severe lung pathology, suggesting a role for sPLA2-X over-expression in inflammatory airway diseases (Curfs, Ghesquiere et al. 2008). The previously identified expression pattern for sPLA2-X does not include cardiac or arterial tissues, suggesting that the enzyme does not play a role in atherosclerosis, however, sPLA2-X is present in blood leukocytes as part of the body's natural defence system and has been shown by immunohistochemistry to exist in atherosclerotic plaques (Curfs, Ghesquiere et al. 2008). However, unlike sPLA2-IIA and V it only appears to be present in circulation rather than localised in normal arteries, suggesting it is more likely to be part of the on-going inflammatory process that perpetuates the progress of an atherosclerotic plaque, rather than contributing to atherosclerotic plaque initiation.

In summary, the data we have collected in this chapter suggests that sPLA2-X, despite its potency, contributes minimally to overall sPLA2 activity in CHD related tissues and is not associated with CHD risk, but further analysis must be carried out in order to assess and verify how representative these data are in larger populations.

CHAPTER 7

Discussion

7.1. General Discussion

The burden of coronary heart disease, as outlined in **Chapter 1**, has become a global problem to overcome as the number one cause of death. This is in spite of significant improvements in medical management (Liu, Maniadakis et al. 2002) and since the incidence of deaths by CVD is predicted to rise to between 6.2 million and 8.2 million by the year 2030 (Mathers and Loncar 2006) the importance of identifying CVD contributing factors to identify potential novel therapies is vital and relevant. **Chapter 1** describes a small portion of the large number of CVD risk factors that have been previously identified. Of these it seems that family history is still considered to be one of the strongest independent risk factors (Lusis, Mar et al. 2004) and this implies that there is a strong cause for continuing to investigate the genetic contribution to CVD. To date, research into the genetics of heart disease has yielded only a small number of variants to explain a minimal portion of disease heritability (Manolio 2010). Additionally these variants add very little to the classical risk stratification techniques (Drenos, Whittaker et al. 2007; Humphries, Drenos et al. 2010). This leaves us with plenty of scope for identifying gene variations associated with genes encoding enzymes that have been previously identified as relevant to CVD. This will help us to understand the true contribution of these enzymes to the disease and, potentially, the heritability of CVD, using the ever improving genomic technologies and novel insights into the 3.1 billion base pairs that make up the human genome. As an example, I present data throughout this thesis to analyse the role of the suspected CVD risk factor, secretory phospholipase A2. In this chapter I hope to summarize the roles of different experimental techniques (**7.1.1- 7.1.4** and **7.1.6**) in medical research and show how Mendelian randomisation (MR) and expression quantitative trait loci (eQTL) analysis could generate relevant data that cannot be conclusively derived from other techniques (**7.11-7.13**), which may be important when developing an inhibitor drug (**7.1.5**). The remainder of the chapter (**7.2**) is an overview of the data presented in the results chapters and a discussion of its value in determining the contribution of 3 sPLA2 isoenzymes to total sPLA2 activity and the potential of each of these enzymes to contribute to CHD, including results from both MR and eQTL analysis.

7.1.1. Observational Data

Observational studies are useful research tools to predict how changes in a particular biomarker could affect a disease outcome. An observational study is an analysis of a group

of individuals that measures relevant biomarkers and observes numbers of disease events in a natural setting without any kind of experimental intervention (Mann 2003; Ligthelm, Borzi et al. 2007). The sample can then be separated into cases and controls dependent on whether an event was observed or not. These two groups can then be compared to see if there is any significant difference in supposed causal biomarkers between the groups, allowing one to conclude whether this biomarker will causally affect the event outcome. Observational studies are often used when it is considered to be unethical to carry out a randomised control trial (RCT) to analyse an association (Ligthelm, Borzi et al. 2007). If it is considered to be unethical to randomly allocate an intervention (or not) to a group of people then an observational study provides an opportunity to study the link in a natural setting. This type of study can also be used in a situation where it is not possible to carry out a blind RCT or if an RCT is impractical. The benefits of this kind of study are that they are inexpensive and there is a much better potential for recruiting high numbers, which can mean a better prediction of rare or adverse events. (Black 1996; Ligthelm, Borzi et al. 2007). The major downside is that observational studies are subject to confounding, for example by another risk factor for the disease (lifestyle or otherwise) and reverse causation, where the true causality of a supposed intervention is masked by its association with another risk factor, or the fact that changes in the supposed intervention are actually as a result of the outcome rather than the other way around. Additionally there is no formal randomisation so there is a higher potential for bias and if the data has not been collected for the purpose of analysing the biomarker of interest it is possible that not all of the relevant measurements have been taken (Ligthelm, Borzi et al. 2007). In this way observational studies may generate false positive results. This can be problematic since drug design and RCT initiation are often based on these results and this is expensive and time consuming.

The observational data described throughout this thesis strongly suggests that sPLA2 could play a causal role for CHD. However, further analysis has shown that this is unlikely and sPLA2s are more likely to be markers of disease, induced by the prolonged inflammatory response associated with CHD. This study highlights the importance of observational data to identify potential biomarkers, but also the need for closer inspection of these relationships before they are taken forward as novel disease intervention pathways.

7.1.2. Animal (*in vivo*) Models

Animal models can be an excellent tool to study the relationship between a suspected risk factor and a disease outcome in a living model. In particular, regarding the sPLA2 effects in CHD, which has been reviewed in this thesis, mouse models of atherosclerosis are the most widely used to study the potential effects of sPLA2 proteins

(Ghesquiere, Gijbels et al. 2005; Getz and Reardon 2012). Mouse models are relatively cheap and easy to keep and feed and their rate of reproduction and litter sizes make it possible to interbreed them and examine differences in particular groups relatively quickly and in large numbers. It is also reasonably easy to genetically manipulate a mouse model to suit the specific purposes of a study (Getz and Reardon 2012) with the creation of knockout, knock-down and transgenic mice. The disadvantages lie in the incongruities between the mouse and human forms of atherosclerosis and sPLA2 distribution. For example; mice do not naturally present with coronary lesions and the lesions that do develop (mostly in the aortic root) are of limited complexity compared to human lesions and distributed differently. The mouse model only represents atherogenesis, not late stage unstable atherosclerosis. There is limited tissue to study, due to the small size of the model and wild-type mice are generally fairly atherosclerosis resistant, in part due to their innate monotypic HDL, increasing their cardio-protectiveness (Getz and Reardon 2012). Specifically considering sPLA2, there is a reasonably strong homology between human and mouse genes (USCS genome browser) (Tischfield, Xia et al. 1996). However, the most frequently used mouse model when researching sPLA2 is the C57/BL6J model, which is a natural knockout model for sPLA2-IIA and has been frequently crossed with known atherosclerosis inducing mouse models such as *Apoe*^{-/-} and *Ldlr*^{-/-} to assess the effects of the lack of sPLA2-IIA in C57/BL6 compared to wild-type models (Ghesquiere, Gijbels et al. 2005; Boyanovsky, Zack et al. 2009). This model has also been useful for breeding a transgenic mouse to express human sPLA2-IIA, which allowed researchers to analyse the impact of human sPLA2-IIA without the overlap of murine sPLA2-IIA confounding the results. It is important to remember that, while mouse models provide excellent background data, which is often comparable to humans, the disease pathway has significant differences and although sPLA2-IIA can be redundant with no obvious side-effects in a natural knock-out mouse model, the same may not be true in humans (Rosengren, Peilot et al. 2006). Additionally the manipulation of genetics and diet within mouse models does not necessarily mimic the natural setting of humans. (Fraser, Hislop et al. 2009).

7.1.3. *In vitro* Analysis

Unlike animal modelling, *in vitro* analysis is not carried out in living organisms. Components of an organism are isolated and analysed using a variety of experimental techniques. The benefit of this is that human tissues and fluids can be used for analysis. Tissue samples taken from relevant organs as well as blood and saliva samples are regularly used to isolate DNA or mRNA for measurement and comparison. In **Chapter 3** DNA extracted from blood samples was used in an array of functional techniques including EMSA, Luciferase assays and gene expression. The ASAP study utilised tissue biopsies to

take measurements of mRNA expression and tissue biopsies are also used for immunohistochemistry. *In vitro* analysis is an excellent way to assess potential biological pathways and associations in a human model. However, *in vitro* work is not carried out in a living and active organism and results from isolated samples may not truly reflect the way things interact in a complex living organism. Additionally, *in vitro* experiments often require the use of vector, which may add to potential confounding.

A strong, combined body of concurrent evidence from *in vitro*, *in vivo* and observational data will provide a solid background on which to make a reasonable conclusion about the likely contribution of a biomarker to disease.

7.1.4. Mendelian Randomisation

Mendelian randomisation (MR) analysis has the ability to be a key bridging step between observational results and an RCT. The concept of MR is simply that a genetic variant that is known to be independently associated with a biomarker of interest can be used to prove the causality of this biomarker in disease (Katan 1986; Davey Smith 2007; Sheehan, Didelez et al. 2008). Alleles are randomly consigned during meiosis and are therefore not subject to reverse causation (Sheehan, Didelez et al. 2008). If a variant is independently associated with naturally different biomarker levels compared to the wild-type it can be considered a candidate for use in MR analysis. Additionally, all other biomarker measurements should be evenly and randomly distributed between carriers of the variant and non-carriers, which will prevent confounding by other risk factors. If large enough numbers can be recruited, we can compare measurements of a disease outcome of interest between carriers and non-carriers. If a significant difference in incidence of disease is detected between these groups we can be sure that it is due to the difference in biomarker levels attributed to the variant and it is therefore likely that the biomarker is causal for disease. If the rate of incidence is not significantly different between the two groups, then the biomarker affected by the variant is not causally associated with disease (Sheehan, Didelez et al. 2008). MR has been previously considered as a natural RCT, where the two randomly assigned groups are carriers of the variant and non-carriers instead of placebo and drug and are compared for differences in outcome in the same way (Hingorani and Humphries 2005). Biomarkers of interest are generally selected based on observational data showing a significant association between the biomarker and a disease outcome (Sheehan, Didelez et al. 2008).

For a genetic variant to be acceptable as an instrument of MR it must fulfil several requirements. First it must be shown to have an independent effect on the biomarker of interest, preferably covering as much of the genetic variance for the biomarker as possible. It

must also be shown not to be associated to any known confounders. Finally, the genetic variant must be independent of disease outcome for known confounders after adjustment (Katan 1986; Sheehan, Didelez et al. 2008).

In brief, MR uses a genetic variant to validate the causal relationship between a biomarker and outcome by examining the relationship of the SNP with both biomarker and outcome and testing the concordance of effects, comparing the observed and expected values (Katan 1986; Sheehan, Didelez et al. 2008).

The limitations of MR can be listed broadly in two categories. Firstly, an inability of a genetic variant to fully meet the criteria set out above for the ideal genetic instrument, which may bias a demonstrated causal effect. For example if the variant only explains a small percentage of the variance of the biomarker. In the case of *PLA2G2A* and rs11573156, the SNP explained ~32% of the variance of *PLA2G2A* mRNA expression and therefore was a very strong instrument. Secondly, extra distributional assumptions are often made when undertaking statistical analysis in MR and these may skew results regarding the magnitude of an effect (though not the presence of the effect itself), for example; considering binary variable outcomes, under typically applied mathematical models, such as logistic regression, conventional linearity is not satisfied. Consequentially, estimations made in the all-linear case should not really be applied to binary outcome data, despite it sometimes being advocated (Sheehan, Didelez et al. 2008).

Despite these limitations, MR is considered to be a reliable and informative analytical technique to add to the profile of data from observational, animal and *in vitro* studies. It is less expensive and less time consuming than drug design and RCT and gives a more accurate indication of the causality of a biomarker than observational data alone. Additionally, with the advent of chip technology where large studies are genotyped for many different SNPs we have large numbers of samples to pool quickly and efficiently. Therefore, it should perhaps be considered a useful practice to carry out an MR analysis before moving on to drug design and RCT, saving valuable time and resources (Hingorani and Humphries 2005). Examples of successful applications of MR analysis are numerous. For instance, regarding heart disease as an outcome, fibrinogen genetic instruments G-455A and C-148T were used to show that fibrinogen association with CHD is probably due to reverse causation or confounding (Smith, Harbord et al. 2005; Keavney, Danesh et al. 2006), but homocysteine variant *MTHFR* C677T showed MR evidence consistent with a causal role for homocysteine in stroke (Casas, Bautista et al. 2005).

7.1.5. Inhibitor Drug Design

Designing molecules to inhibit pathways related to biomarkers known to be causal for a disease is an important step to potentially preventing a disease outcome. Once designed, drugs undergo extensive efficacy and safety testing, which start (as we have seen above) with *in vitro* data and animal model testing. The supposed gold standard test of a drug is a large randomised control trial (RCT) in humans. Many inhibitor molecules are designed based on data from observational studies alone and tested *in vitro* and in animal models, which may yield positive results regarding the inhibition of the causal biomarker and decreased incidence of disease. However, these molecules sometimes generate disappointing results when used in large scale RCTs and, at times, the premature cessation of an RCT has happened due to increased undesirable side effects or even mortality seen with the drug. When a drug trial is terminated early for these reasons it is important to understand whether it is due to an off target effect of the molecule itself, suggesting that a different class of drugs working on the same biomarker may not have the same side-effects and remain a useful novel therapeutic molecule, or if the inhibition of the pathway is the cause of the side-effect, in which case the inhibition of this biomarker can no longer be considered as a novel intervention since any drug designed to inhibit this pathway would cause the same side-effect. A lot of time and resources could be saved if we attempted to answer these questions before they arise and MR is an excellent tool to predict whether a change in biomarker levels is naturally associated with any undesirable side-effects. There are several examples where MR has been useful in this way. As described in **Chapter 1.7.3**.

7.1.6. ASAP Study: The Value of Expression Quantitative Trait Loci (eQTLs)

Expression quantitative trait loci (eQTLs) are defined as genomic loci that are associated with the regulation of gene transcription or gene expression. Static eQTLs are considered to remain constant across all tissues and dynamic eQTLs are tissue dependent. An eQTL can also be *cis* (occurring and acting on the gene of interest) or *trans* (acting on a proximal gene, often on the same chromosome, (Damerval, Maurice et al. 1994; Petretto, Mangion et al. 2006)).

Results from the ASAP study (Folkersen, van't Hooft et al. 2010) were presented in this thesis. The gene and SNP specific eQTL data measured mRNA and gave us excellent evidence to substantiate our findings about the interactions and contributions of each sPLA2 isoform to sPLA2 activity. We could also use these eQTLs as additional proof that the sPLA2-IIA functional SNP that we selected as a genetic instrument for MR in **Chapter 3** and **Chapter 4** was truly regulatory for mRNA expression.

7.2. Overall Summary of Results

Phospholipases including secretory phospholipase had been observed to be significantly associated with CHD, specifically atherosclerosis. Common genetic variants for the genes that code for these enzymes can be used for techniques such as Mendelian randomisation in order to determine whether an increase in secretory phospholipases contributes causally to the onset of atherosclerosis, or is a result of reverse causation or confounding and thus a marker of disease. In this thesis one of the aims was to identify common genetic variants to genotype in numerous prospective and case control studies. These studies were in both general populations and acute coronary syndrome settings. The results from these studies were used in order to improve the current understanding of the role of secretory phospholipases in the development of CHD and clarify the suitability of sPLA2s as potential therapeutic targets. It was also an aim to investigate the genetics of sPLA2-IIA, V and X in order to improve the understanding of how these enzymes interact with and compensate for each other as well as the individual contributions of each enzyme to overall sPLA2 activity.

Refining these aims further, the main focus of this thesis has been to try and identify the true role of sPLA2-IIA in the risk of CHD through the analysis of common variants within the sPLA2-IIA coding gene, *PLA2G2A*. However, a small follow-up section has been dedicated to the potential contributions of sPLA2-V and sPLA2-X to sPLA2 activity and CHD risk.

I believe that I successfully identified two truly functional variants within *PLA2G2A*, rs11573156 C>G and rs3767221 T>G and used rs11573156 as a genetic instrument to show conclusively in a Mendelian randomisation of over 30 studies that sPLA2-IIA is not causal for CHD, but merely a marker of the disease. Assessing the contributions of each enzyme to sPLA2 activity was carried out by studying the expression analysis of each gene in different tissues in the ASAP study and the results of several meta-analyses along with previously published data regarding the potential differential contributions of each SNP. I think several conclusions can be drawn from this combination of data. The contribution of sPLA2 isoenzymes to total sPLA2 activity appears to be tissue and environment dependent (i.e. where the tissue is situated). Previous data has shown that sPLA2-X is most highly expressed in thymus, spleen and peripheral blood leukocytes (Curfs, Ghesquiere et al. 2008). Data from the ASAP study that shows that SNPs were associated with consistently low levels of *PLA2G10* mRNA across all the tissues measured in the study, which did not include thymus, spleen or blood leukocytes. Figure 7.1 shows the overall mRNA expression for each gene probeset, which consists of a set of probes designed to measure the mRNA

expression of each exon of each gene (*PLA2G2A*, *PLA2G5* and *PLA2G10*) across all the tissues of the study. The results show that the probeset with the highest mRNA expression across these tissues was *PLA2G2A* and then *PLA2G5*. The probeset covering *PLA2G10* mRNA expression was so low across these tissues that it was barely measurable. It should be noted that a slight flaw of this study was the low SNP coverage for *PLA2G10* compared to *PLA2G2A* and *PLA2G5*, which could potentially confound the mRNA expression results.

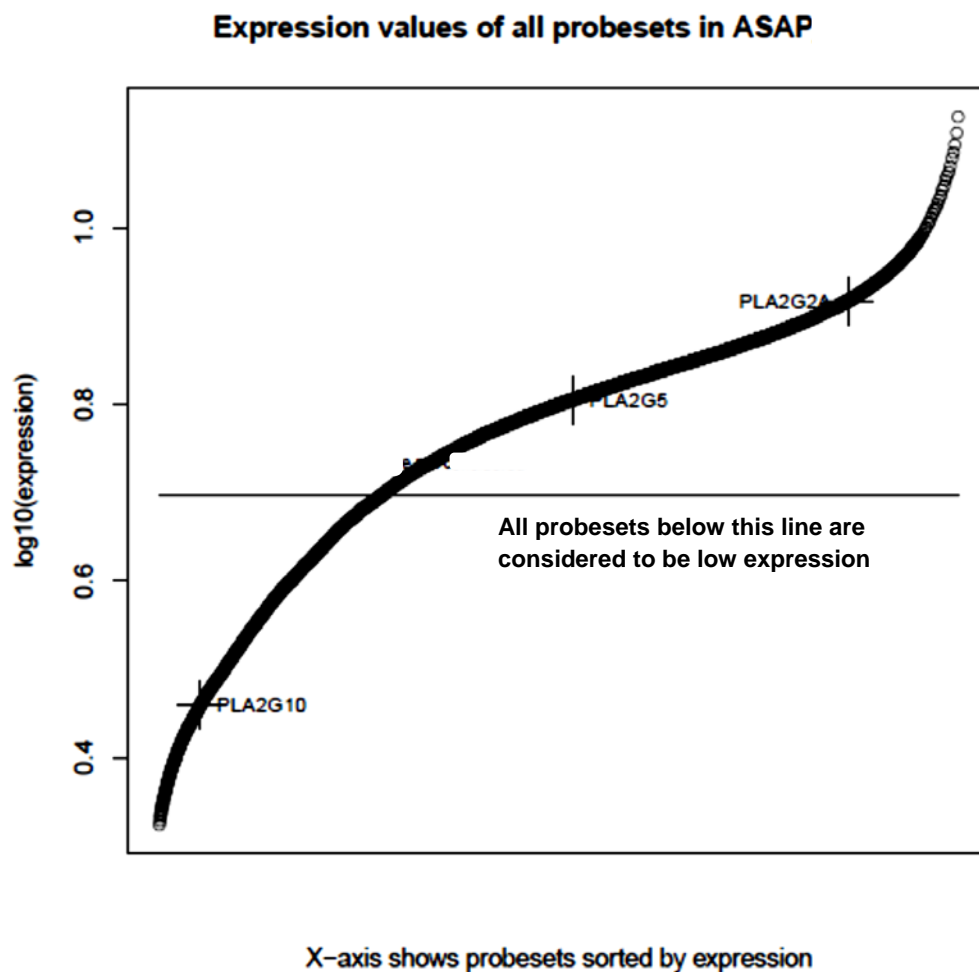


Figure 7.1 Shows the *PLA2* mRNA relative expression values across all *PLA2* isoform probesets used in the ASAP study. *PLA2G2A* covers 101 SNPs, *PLA2G5* covers 123 SNPs and *PLA2G10* covers 13 SNPs.

Despite the low SNP coverage for *PLA2G10*, these results are further validated by previous data that has shown *PLA2G2A* and *PLA2G5* to be most highly expressed in several of the tissues covered by ASAP (Crowl, Stoller et al. 1991; Murakami and Kudo 1999;

Masuda, Murakami et al. 2005), suggesting that *PLA2G10* mRNA is truly the least expressed in the tissues across this study. Considering the results from the meta-analysis studies the SNPs associated with sPLA2 activity levels were localized to the *PLA2G2A* gene. This suggests that sPLA2-IIA has the strongest contributions to sPLA2 activity, with the *PLA2G2A* SNP rs11573156 showing the strongest overall association. The *PLA2G5* SNP rs525380 showed the strongest association with *PLA2G5* mRNA suggesting that it should have a strong association with sPLA2-V activity. The lack of a significant association between this SNP and overall sPLA2 activity therefore suggests that sPLA2-V does not contribute as significantly to overall sPLA2 activity compared to sPLA2-IIA. None of the *PLA2G10* SNPs were significantly associated with sPLA2 activity. (Curfs, Ghesquiere et al. 2008; Lambeau and Gelb 2008). The IC₅₀ values for each enzyme considering the drug varespladib in humans were quoted as 9nM for sPLA2-IIA, 15nM for sPLA2-X and 77nM for sPLA2-V, which explains the ability of the drug to significantly lower sPLA2-IIA levels and to lower sPLA2 activity to immeasurably low levels (Fraser, Hislop et al. 2009). These values fit nicely with the rank order potency amongst human sPLA2 to kill Gram-positive bacteria *in vitro*; sPLA2-IIA > sPLA2-X > sPLA2-V (Lambeau and Gelb 2008). It seems, therefore, that in circulation the contribution of each enzyme to activity is likely to be in the same order; sPLA2-IIA > sPLA2-X > sPLA2-V, but in tissues local to the heart and in the liver it seems more likely to be sPLA2-IIA > sPLA2-V > sPLA2-X, with sPLA2-IIA and sPLA2-V potentially having a confounding effect on the magnitude of each enzymes individual effect. The increased levels of sPLA2-IIA and its strong proteoglycan binding ability (Mallat, Lambeau et al. 2010) appear to give this enzyme an increased contribution to sPLA2 activity compared to the high potency of sPLA2-X (Lambeau and Gelb 2008). Considering sPLA2-V, this enzyme appears to contribute less to circulating sPLA2 activity levels, but is likely to be expressed more highly in specific tissues. Again, it is important to recognise the possibility that sPLA2-IIA and sPLA2-V expression are co-regulated and inherited, and this may be confounding. The specific expression and localization of these two enzymes is different, perhaps because only one of these enzymes is truly required in any given tissue based on the slight differences between them that have been described previously. Therefore, should there be a deficiency in the regular enzyme in that tissue it may be that a certain level of compensation by the other enzyme could occur. The fact that these two genes are closely linked (Tischfield, Xia et al. 1996) may therefore make sPLA2-V less important in contributing to circulating sPLA2-activity, but more highly expressed and therefore potentially important in, for example, the tissues and cells of human airways (Masuda, Murakami et al. 2005). Data from ASAP and previous tissue expression data suggest that, regarding the onset of atherosclerosis, the two enzymes most likely to contribute to sPLA2-activity in CHD related tissues are sPLA2-IIA and sPLA2-V as these enzymes are found in normal arteries and cardiac tissue as well as in

circulation (Crowl, Stoller et al. 1991; Murakami and Kudo 1999; Masuda, Murakami et al. 2005). Since the gene *PLA2G2A* appears to have the greatest contribution to sPLA2 activity, sPLA2-IIA would be the most essential enzyme to inhibit if sPLA2 was causally linked to CHD. However, as concluded in **Chapter 4**, this enzyme is highly unlikely to contribute the onset of CHD, suggesting that none of these *PLA2* isoenzymes is likely to be causal regarding primary CHD. When considering secondary CHD, past evidence has suggested that all 3 enzymes have been detected in atherosclerotic lesions when over-expressed in transgenic models (Hurt-Camejo, Andersen et al. 1997; Karabina, Brocheriou et al. 2006; Bostrom, Boyanovsky et al. 2007) and in human atherosclerotic lesions (Elinder, Dumitrescu et al. 1997; Sartipy, Johansen et al. 2000; Karabina, Brocheriou et al. 2006; Rosengren, Peilot et al. 2006; Kimura-Matsumoto, Ishikawa et al. 2008) and this suggests that sPLA2 enzymes could be part of the perpetual inflammatory reaction associated with atherosclerosis, contributing to driving the upregulation of further inflammatory enzymes and thus becoming a marker of atherosclerosis. However, I feel that further genetic investigation of SNPs related to *PLA2G5* and *PLA2G10* is necessary in the case of secondary events. Recently Anthera terminated the Phase III trial of varespladib, an sPLA2 inhibitor, early due to concerns about lack of efficacy (<http://investor.anthera.com/releasedetail.cfm?ReleaseID=656340>). The drug is reported to inhibit all three isoforms, sPLA2-IIA, V and X, and a lack of efficacy of this inhibitory drug adds to the suggestion that the presence of these enzymes in atherosclerotic lesions is merely as a marker and not associated causally to secondary cardiac events (Anthera, 2012). The trial was due to recruit 6,500 high-risk ACS patients across 15 countries in the USA and Europe and was to test a combination therapy of varespladib methyl (500mg once daily) with atorvastatin. The primary end-point was the effect on major adverse coronary events such as cardiovascular death, unstable angina, myocardial infarction or stroke (Nicholls, Cavender et al. 2011). Our large scale MR analysis presented in **Chapter 4** of this thesis was conducted at the same time and results suggest that sPLA2 falls into the same category as CRP, where the independent effect seen in observational studies and early positive *in vitro* and *in vivo* results (Snyder, Bach et al. 1999; Fraser, Hislop et al. 2009) for the drug are not reflected in the MR study, which shows no significant association between increased sPLA2-IIA and incidence of CHD. Based on the MR analysis it seems likely that sPLA2-IIA inhibition will not decrease CHD incidence, regardless of the class of molecule used. Had this data been available prior to sPLA2 inhibitor design, perhaps more robust initial testing of the sPLA2/CHD relationship would have been undertaken, saving time and resources for the pharmaceutical industry. I feel that the work in **Chapter 4** shows how Mendelian randomisation could be used to predict causality and potentially be of significant practical use to drug companies in this way.

7.2.1. Limitations

Like all scientific research this thesis encountered limitations and several are of note. The assay used to measure sPLA2 activity in all of the observational studies dealt with in this thesis is inclusive of all three sPLA2 isoforms; sPLA2-IIA, sPLA2-V and sPLA2-X. At this time there are no reliable assays able to measure each individual isoform (Mallat, Steg et al. 2005). Additionally there is an assay for sPLA2-IIA levels, but there is no ELISA available to measure either sPLA2-V or -X (Mallat, Steg et al. 2005). This makes it incredibly difficult to assess the contributions of each enzyme individually and ultimately determine whether any one enzyme is more likely to contribute to total sPLA2 activity or CHD.

There was significant difficulty genotyping the *PLA2G10* SNPs rs4003228 and rs4003232. Originally I attempted to genotype these SNPs using TaqMan genotyping specific assays (40x), but these assays failed to work. Standard RFLP techniques were then employed to genotype these SNPs in several studies, but unfortunately accurate genotyping was still unattainable. Eventually I carried out a nested PCR prior to RFLP, which was very time consuming and led to less studies being genotyped and thus a lower powered meta-analysis including these SNPs.

The majority of the SNPs assessed throughout this thesis were identified as SNPs of interest or tagging-SNPs in previously published studies. It is possible that the SNP identification techniques used in these studies have now become outdated or that new SNPs have been identified and genotyped since their publication and this suggests that we cannot be entirely sure that these are the best candidate SNPs for investigation. However, the addition of the eQTL data from the ASAP study has taken away some of this uncertainty and allowed us to validate several of the SNPs as being potentially functional.

Regarding the EMSA analysis in **Chapter 3** relating to rs3767221 A>C, I was unable to produce a supershift EMSA due to time constraints, which would have distinguished the protein binding differentially in the region of the rare C allele and proved specific, differential binding occurred here. This would also give a more significant insight into the potential biological mechanism at work here.

7.3. Conclusions and Further Work

In conclusion, based on the results documented in this thesis, I think that it is likely that sPLA2-IIA has the biggest contribution to circulating sPLA2 activity, though the rank order of contributions from each enzyme is likely to change in a tissue dependent manner perhaps due to the differing mechanisms of each enzyme producing variable activity in different tissues. The genes encoding sPLA2-IIA and sPLA2-V may be partially co-regulated

and display high linkage. It is therefore possible that these enzymes have a certain amount of overlap and compensatory behaviour for each other. Regarding CHD, a comprehensive Mendelian randomisation analysis of sPLA2-IIA using the functional SNP rs11573156 showed that this enzyme was unlikely to be causal for CHD. Since sPLA2-X is not expressed in normal arteries and both sPLA2-V and sPLA2-X appear to contribute less to circulatory sPLA2 activity it is also unlikely that these enzymes are individually associated with the onset of CHD.

In order to confirm these conclusions, further analysis should be carried out. Functionality of SNPs for sPLA2-V and sPLA2-X should be confirmed and then genotyped in significantly powered meta-analyses so that Mendelian randomisation analyses can be carried out and compared to that of sPLA2-IIA and examined regarding CHD/CHD related disorders. It may also be important to further assess the co-regulation between the genes of the *PLA2* cluster on chromosome 1 to further explore the impact of this on each individual enzyme and the overall actions of the cluster. Additionally, several other sPLA2 genes occur outside the chromosome 1 cluster, like *PLA2G10* and it would be useful to find out more about these. To this end, investigations into the genes *PLA2G3* and *PLA21B* are already underway.

Bibliography

- (2001). "Rationale and design of the GRACE (Global Registry of Acute Coronary Events) Project: a multinational registry of patients hospitalized with acute coronary syndromes." Am Heart J 141(2): 190-199.
- Ackermann, E. J., E. S. Kempner, et al. (1994). "Ca(2+)-independent cytosolic phospholipase A2 from macrophage-like P388D1 cells. Isolation and characterization." J Biol Chem 269(12): 9227-9233.
- Afonina, I., M. Zivarts, et al. (1997). "Efficient priming of PCR with short oligonucleotides conjugated to a minor groove binder." Nucleic Acids Res 25(13): 2657-2660.
- Ali, M. K., K. M. Narayan, et al. (2010). "Diabetes & coronary heart disease: current perspectives." Indian J Med Res 132(5): 584-597.
- Ambrose, J. A. and R. S. Barua (2004). "The pathophysiology of cigarette smoking and cardiovascular disease: an update." J Am Coll Cardiol 43(10): 1731-1737.
- Andreotti, F., I. Porto, et al. (2002). "Inflammatory gene polymorphisms and ischaemic heart disease: review of population association studies." Heart 87(2): 107-112.
- Antonio, V., A. Brouillet, et al. (2002). "Transcriptional regulation of the rat type IIA phospholipase A2 gene by cAMP and interleukin-1beta in vascular smooth muscle cells: interplay of the CCAAT/enhancer binding protein (C/EBP), nuclear factor-kappaB and Ets transcription factors." Biochem J 368(Pt 2): 415-424.
- Asmis, R., C. Randriamampita, et al. (1994). "Intracellular Ca²⁺, inositol 1,4,5-trisphosphate and additional signalling in the stimulation by platelet-activating factor of prostaglandin E2 formation in P388D1 macrophage-like cells." Biochem J 298 Pt 3: 543-551.
- Assmann, G., H. Schulte, et al. (1996). "High-density lipoprotein cholesterol as a predictor of coronary heart disease risk. The PROCAM experience and pathophysiological implications for reverse cholesterol transport." Atherosclerosis 124 Suppl: S11-20.
- Aulchenko, Y. S., S. Ripatti, et al. (2009). "Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts." Nat Genet 41(1): 47-55.
- Badimon, L., R. F. Storey, et al. (2011). "Update on lipids, inflammation and atherothrombosis." Thromb Haemost 105 Suppl 1: S34-42.
- Baldassarre, D., K. Nyssönen, et al. (2010). "Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study." Eur Heart J 31(5): 614-622.
- Balestrieri, B., V. W. Hsu, et al. (2006). "Group V secretory phospholipase A2 translocates to the phagosome after zymosan stimulation of mouse peritoneal macrophages and regulates phagocytosis." J Biol Chem 281(10): 6691-6698.
- Balsinde, J., S. E. Barbour, et al. (1994). "Arachidonic acid mobilization in P388D1 macrophages is controlled by two distinct Ca(2+)-dependent phospholipase A2 enzymes." Proc Natl Acad Sci U S A 91(23): 11060-11064.
- Balsinde, J., H. Shinohara, et al. (1999). "Group V phospholipase A(2)-dependent induction of cyclooxygenase-2 in macrophages." J Biol Chem 274(37): 25967-25970.
- Barter, P. J., M. Caulfield, et al. (2007). "Effects of torcetrapib in patients at high risk for coronary events." N Engl J Med 357(21): 2109-2122.
- Barter, P. J., S. Nicholls, et al. (2004). "Antiinflammatory properties of HDL." Circ Res 95(8): 764-772.
- Beers, S. A., A. G. Buckland, et al. (2003). "Effect of tryptophan insertions on the properties of the human group IIA phospholipase A2: mutagenesis produces an enzyme with characteristics similar to those of the human group V phospholipase A2." Biochemistry 42(24): 7326-7338.

- Beisiegel, U., W. Weber, et al. (1989). "The LDL-receptor-related protein, LRP, is an apolipoprotein E-binding protein." *Nature* 341(6238): 162-164.
- Berliner, J. A. and J. W. Heinecke (1996). "The role of oxidized lipoproteins in atherogenesis." *Free Radic Biol Med* 20(5): 707-727.
- Berra, K. (2000). "Women, coronary heart disease, and dyslipidemia: does gender alter detection, evaluation, or therapy?" *J Cardiovasc Nurs* 14(2): 59-78.
- Beulens, J. W., E. M. Monninkhof, et al. (2010). "Cohort profile: the EPIC-NL study." *Int J Epidemiol* 39(5): 1170-1178.
- Bezzine, S., J. G. Bollinger, et al. (2002). "On the binding preference of human groups IIA and X phospholipases A2 for membranes with anionic phospholipids." *J Biol Chem* 277(50): 48523-48534.
- Bezzine, S., R. S. Koduri, et al. (2000). "Exogenously added human group X secreted phospholipase A(2) but not the group IB, IIA, and V enzymes efficiently release arachidonic acid from adherent mammalian cells." *J Biol Chem* 275(5): 3179-3191.
- Bhupathiraju, S. N. and K. L. Tucker (2011). "Coronary heart disease prevention: nutrients, foods, and dietary patterns." *Clin Chim Acta* 412(17-18): 1493-1514.
- Birts, C. N., C. H. Barton, et al. (2010). "Catalytic and non-catalytic functions of human IIA phospholipase A2." *Trends Biochem Sci* 35(1): 28-35.
- Birts, C. N., L. M. Bergman, et al. (2011). "CtBPs promote mitotic fidelity through their activities in the cell nucleus." *Oncogene* 30(11): 1272-1280.
- Birts, C. N., R. Harding, et al. (2010). "Expression of CtBP family protein isoforms in breast cancer and their role in chemoresistance." *Biol Cell* 103(1): 1-19.
- Blache, D., T. Gautier, et al. (2012). "Activated platelets contribute to oxidized low-density lipoproteins and dysfunctional high-density lipoproteins through a phospholipase A2-dependent mechanism." *FASEB J* 26(2): 927-937.
- Black, N. (1996). "Why we need observational studies to evaluate the effectiveness of health care." *BMJ* 312(7040): 1215-1218.
- Blake, G. J. and P. M. Ridker (2003). "C-reactive protein and other inflammatory risk markers in acute coronary syndromes." *J Am Coll Cardiol* 41(4 Suppl S): 37S-42S.
- Boekholdt, S. M., T. T. Keller, et al. (2005). "Serum levels of type II secretory phospholipase A2 and the risk of future coronary artery disease in apparently healthy men and women: the EPIC-Norfolk Prospective Population Study." *Arterioscler Thromb Vasc Biol* 25(4): 839-846.
- Boer, J. M., E. J. Feskens, et al. (1999). "The joint impact of family history of myocardial infarction and other risk factors on 12-year coronary heart disease mortality." *Epidemiology* 10(6): 767-770.
- Boilard, E., Y. Lai, et al. (2010). "A novel anti-inflammatory role for secretory phospholipase A2 in immune complex-mediated arthritis." *EMBO Mol Med* 2(5): 172-187.
- Borissoff, J. I., H. M. Spronk, et al. (2011). "The hemostatic system as a modulator of atherosclerosis." *N Engl J Med* 364(18): 1746-1760.
- Bostrom, M. A., B. B. Boyanovsky, et al. (2007). "Group v secretory phospholipase A2 promotes atherosclerosis: evidence from genetically altered mice." *Arterioscler Thromb Vasc Biol* 27(3): 600-606.
- Boyanovsky, B., M. Zack, et al. (2009). "The capacity of group V sPLA2 to increase atherogenicity of ApoE-/- and LDLR-/- mouse LDL in vitro predicts its atherogenic role in vivo." *Arterioscler Thromb Vasc Biol* 29(4): 532-538.
- Boyanovsky, B. B., D. R. van der Westhuyzen, et al. (2005). "Group V secretory phospholipase A2-modified low density lipoprotein promotes foam cell formation by a SR-A- and CD36-independent process that involves cellular proteoglycans." *J Biol Chem* 280(38): 32746-32752.
- Brown, K. E., S. S. Guest, et al. (1997). "Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin." *Cell* 91(6): 845-854.

- Brown, M. S. and J. L. Goldstein (1986). "A receptor-mediated pathway for cholesterol homeostasis." Science 232(4746): 34-47.
- Brown, M. S. and J. L. Goldstein (1997). "The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor." Cell 89(3): 331-340.
- Buckland, A. G. and D. C. Wilton (2000). "The antibacterial properties of secreted phospholipases A(2)." Biochim Biophys Acta 1488(1-2): 71-82.
- Burgess, S., S. Seaman, et al. (2011). "Missing data methods in Mendelian randomization studies with multiple instruments." Am J Epidemiol 174(9): 1069-1076.
- Burke, J. E. and E. A. Dennis (2009). "Phospholipase A2 structure/function, mechanism, and signaling." J Lipid Res 50 Suppl: S237-242.
- Burkett, K. M., M. Ghadessi, et al. (2005). "A comparison of five methods for selecting tagging single-nucleotide polymorphisms." BMC Genet 6 Suppl 1: S71.
- Businaro, R., A. Tagliani, et al. (2012). "Cellular and molecular players in the atherosclerotic plaque progression." Ann N Y Acad Sci 1262: 134-141.
- Calabrese, C., M. Triggiani, et al. (2000). "Arachidonic acid metabolism in inflammatory cells of patients with bronchial asthma." Allergy 55 Suppl 61: 27-30.
- Cambou, J. P., T. Simon, et al. (2007). "The French registry of Acute ST elevation or non-ST-elevation Myocardial Infarction (FAST-MI): study design and baseline characteristics." Arch Mal Coeur Vaiss 100(6-7): 524-534.
- Cannon, C. P., E. Braunwald, et al. (2004). "Intensive versus moderate lipid lowering with statins after acute coronary syndromes." N Engl J Med 350(15): 1495-1504.
- Cannon, R. O., 3rd (1998). "Role of nitric oxide in cardiovascular disease: focus on the endothelium." Clin Chem 44(8 Pt 2): 1809-1819.
- Carew, T. E., D. C. Schwenke, et al. (1987). "Antiatherogenic effect of probucol unrelated to its hypocholesterolemic effect: evidence that antioxidants in vivo can selectively inhibit low density lipoprotein degradation in macrophage-rich fatty streaks and slow the progression of atherosclerosis in the Watanabe heritable hyperlipidemic rabbit." Proc Natl Acad Sci U S A 84(21): 7725-7729.
- Carmena, R., P. Duriez, et al. (2004). "Atherogenic lipoprotein particles in atherosclerosis." Circulation 109(23 Suppl 1): III2-7.
- Casas, J. P., L. E. Bautista, et al. (2005). "Homocysteine and stroke: evidence on a causal link from mendelian randomisation." Lancet 365(9455): 224-232.
- Casas, J. P., E. Ninio, et al. (2010). "PLA2G7 genotype, lipoprotein-associated phospholipase A2 activity, and coronary heart disease risk in 10 494 cases and 15 624 controls of European Ancestry." Circulation 121(21): 2284-2293.
- Chatterjee, S. and J. K. Pal (2009). "Role of 5'- and 3'-untranslated regions of mRNAs in human diseases." Biol Cell 101(5): 251-262.
- Chaudhury, A., G. S. Hussey, et al. (2011). "3'-UTR-mediated post-transcriptional regulation of cancer metastasis: beginning at the end." RNA Biol 8(4): 595-599.
- Chen, B.-Y. and H. W. Janes (2002). PCR cloning protocols. Totowa, N.J., Humana Press.
- Chobanian, A. V., G. L. Bakris, et al. (2003). "Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure." Hypertension 42(6): 1206-1252.
- Collaboration, C. C. G. (2008). "Collaborative pooled analysis of data on C-reactive protein gene variants and coronary disease: judging causality by Mendelian randomisation." Eur J Epidemiol 23(8): 531-540.
- Collaboration, C. R. P. C. H. D. G., F. Wensley, et al. (2011). "Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data." BMJ 342: d548.
- Consortium, E. P., B. E. Bernstein, et al. (2012). "An integrated encyclopedia of DNA elements in the human genome." Nature 489(7414): 57-74.

- Cooney, M. T., A. Dudina, et al. (2009). "HDL cholesterol protects against cardiovascular disease in both genders, at all ages and at all levels of risk." Atherosclerosis 206(2): 611-616.
- Corson, M. A. (2009). "Emerging inflammatory markers for assessing coronary heart disease risk." Curr Cardiol Rep 11(6): 452-459.
- Crabb, B. S., B. M. Cooke, et al. (1997). "Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress." Cell 89(2): 287-296.
- Crowl, R. M., T. J. Stoller, et al. (1991). "Induction of phospholipase A2 gene expression in human hepatoma cells by mediators of the acute phase response." J Biol Chem 266(4): 2647-2651.
- Cupillard, L., K. Koumanov, et al. (1997). "Cloning, chromosomal mapping, and expression of a novel human secretory phospholipase A2." J Biol Chem 272(25): 15745-15752.
- Curfs, D. M., S. A. Ghesquiere, et al. (2008). "Macrophage secretory phospholipase A2 group X enhances anti-inflammatory responses, promotes lipid accumulation, and contributes to aberrant lung pathology." J Biol Chem 283(31): 21640-21648.
- Damerval, C., A. Maurice, et al. (1994). "Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of genome expression." Genetics 137(1): 289-301.
- Davey Smith, G. (2007). "Capitalizing on Mendelian randomization to assess the effects of treatments." J R Soc Med 100(9): 432-435.
- Davey Smith, G. and S. Ebrahim (2003). "'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?" Int J Epidemiol 32(1): 1-22.
- Davies, M. J. and N. Woolf (1993). "Atherosclerosis: what is it and why does it occur?" Br Heart J 69(1 Suppl): S3-11.
- Day, N., S. Oakes, et al. (1999). "EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer." Br J Cancer 80 Suppl 1: 95-103.
- De Taeye, B., L. H. Smith, et al. (2005). "Plasminogen activator inhibitor-1: a common denominator in obesity, diabetes and cardiovascular disease." Curr Opin Pharmacol 5(2): 149-154.
- Deconinck, A. E., J. A. Rafael, et al. (1997). "Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy." Cell 90(4): 717-727.
- Dehghan, A., J. Dupuis, et al. (2011). "Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels." Circulation 123(7): 731-738.
- Dennis, E. A. (1994). "Diversity of group types, regulation, and function of phospholipase A2." J Biol Chem 269(18): 13057-13060.
- Deregnacourt, C. and J. Schrevel (2000). "Bee venom phospholipase A2 induces stage-specific growth arrest of the intraerythrocytic Plasmodium falciparum via modifications of human serum components." J Biol Chem 275(51): 39973-39980.
- Dhamrait, S. S., J. R. Payne, et al. (2003). "Variation in bradykinin receptor genes increases the cardiovascular risk associated with hypertension." Eur Heart J 24(18): 1672-1680.
- Dhamrait, S. S., J. W. Stephens, et al. (2004). "Cardiovascular risk in healthy men and markers of oxidative stress in diabetic men are associated with common variation in the gene for uncoupling protein 2." Eur Heart J 25(6): 468-475.
- Diercks, G. F., W. M. Janssen, et al. (2000). "Rationale, design, and baseline characteristics of a trial of prevention of cardiovascular and renal disease with fosinopril and pravastatin in nonhypertensive, nonhypercholesterolemic subjects with microalbuminuria (the Prevention of RENal and Vascular ENdstage Disease Intervention Trial [PREVEND IT])." Am J Cardiol 86(6): 635-638.
- Doll, R. and A. B. Hill (1966). "Mortality of British doctors in relation to smoking: observations on coronary thrombosis." Natl Cancer Inst Monogr 19: 205-268.
- Dorresteijn, J. A., F. L. Visseren, et al. (2011). "Estimating treatment effects for individual patients based on the results of randomised clinical trials." BMJ 343: d5888.

- Drenos, F., J. C. Whittaker, et al. (2007). "The use of meta-analysis risk estimates for candidate genes in combination to predict coronary heart disease risk." Ann Hum Genet 71(Pt 5): 611-619.
- Duscha, B. D., C. A. Slentz, et al. (2005). "Effects of exercise training amount and intensity on peak oxygen consumption in middle-age men and women at risk for cardiovascular disease." Chest 128(4): 2788-2793.
- Dutour, A., V. Achard, et al. (2010). "Secretory type II phospholipase A2 is produced and secreted by epicardial adipose tissue and overexpressed in patients with coronary artery disease." J Clin Endocrinol Metab 95(2): 963-967.
- Dzavik, V., S. Lavi, et al. (2010). "The sPLA2 Inhibition to Decrease Enzyme Release after Percutaneous Coronary Intervention (SPIDER-PCI) trial." Circulation 122(23): 2411-2418.
- Eckey, R., M. Menschikowski, et al. (1997). "Minimal oxidation and storage of low density lipoproteins result in an increased susceptibility to phospholipid hydrolysis by phospholipase A2." Atherosclerosis 132(2): 165-176.
- Eisenberg, S. (1984). "High density lipoprotein metabolism." J Lipid Res 25(10): 1017-1058.
- Eisenberg, S. and R. I. Levy (1975). "Lipoprotein metabolism." Adv Lipid Res 13: 1-89.
- Elinder, L. S., A. Dumitrescu, et al. (1997). "Expression of phospholipase A2 isoforms in human normal and atherosclerotic arterial wall." Arterioscler Thromb Vasc Biol 17(10): 2257-2263.
- Elliott, P., J. C. Chambers, et al. (2009). "Genetic Loci associated with C-reactive protein levels and risk of coronary heart disease." JAMA 302(1): 37-48.
- Erhardt, L. (2009). "Cigarette smoking: an undertreated risk factor for cardiovascular disease." Atherosclerosis 205(1): 23-32.
- Escoffier, J., I. Jemel, et al. (2010). "Group X phospholipase A2 is released during sperm acrosome reaction and controls fertility outcome in mice." J Clin Invest 120(5): 1415-1428.
- Fan, Q., M. Paradon, et al. (1997). "C/EBP factor suppression of inhibition of type II secreted phospholipase A2 promoter in HepG2 cells: possible role of single-strand binding proteins." Mol Cell Biol 17(8): 4238-4248.
- Fenard, D., G. Lambeau, et al. (2001). "A peptide derived from bee venom-secreted phospholipase A2 inhibits replication of T-cell tropic HIV-1 strains via interaction with the CXCR4 chemokine receptor." Mol Pharmacol 60(2): 341-347.
- Fielding, C. J. and P. E. Fielding (1995). "Molecular physiology of reverse cholesterol transport." J Lipid Res 36(2): 211-228.
- Flood, C., M. Gustafsson, et al. (2004). "Molecular mechanism for changes in proteoglycan binding on compositional changes of the core and the surface of low-density lipoprotein-containing human apolipoprotein B100." Arterioscler Thromb Vasc Biol 24(3): 564-570.
- Fodor, J. G., R. Tzerovska, et al. (2004). "Do we diagnose and treat coronary heart disease differently in men and women?" Wien Med Wochenschr 154(17-18): 423-425.
- Folkersen, L., D. Diez, et al. (2009). "GeneRegionScan: a Bioconductor package for probe-level analysis of specific, small regions of the genome." Bioinformatics 25(15): 1978-1979.
- Folkersen, L., F. van't Hooft, et al. (2010). "Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease." Circ Cardiovasc Genet 3(4): 365-373.
- Fowkes, F. G., E. Housley, et al. (1991). "Edinburgh Artery Study: prevalence of asymptomatic and symptomatic peripheral arterial disease in the general population." Int J Epidemiol 20(2): 384-392.
- Fox, C. S., M. G. Larson, et al. (2005). "Glycemic status and development of kidney disease: the Framingham Heart Study." Diabetes Care 28(10): 2436-2440.
- Fox, C. S., M. G. Larson, et al. (2006). "Cross-sectional association of kidney function with valvular and annular calcification: the Framingham heart study." J Am Soc Nephrol 17(2): 521-527.
- Franklin, S. S., J. R. Pio, et al. (2005). "Predictors of new-onset diastolic and systolic hypertension: the Framingham Heart Study." Circulation 111(9): 1121-1127.

- Fraser, H., C. Hislop, et al. (2009). "Varespladib (A-002), a secretory phospholipase A2 inhibitor, reduces atherosclerosis and aneurysm formation in ApoE^{-/-} mice." J Cardiovasc Pharmacol 53(1): 60-65.
- Freedman, D. S., J. D. Otvos, et al. (1998). "Relation of lipoprotein subclasses as measured by proton nuclear magnetic resonance spectroscopy to coronary artery disease." Arterioscler Thromb Vasc Biol 18(7): 1046-1053.
- Furue, S., K. Mikawa, et al. (2001). "Therapeutic time-window of a group IIA phospholipase A2 inhibitor in rabbit acute lung injury: correlation with lung surfactant protection." Crit Care Med 29(4): 719-727.
- Garcia-Garcia, H. M. and P. W. Serruys (2009). "Phospholipase A2 inhibitors." Curr Opin Lipidol 20(4): 327-332.
- Garcia, M. J., P. M. McNamara, et al. (1974). "Morbidity and mortality in diabetics in the Framingham population. Sixteen year follow-up study." Diabetes 23(2): 105-111.
- Getz, G. S. and C. A. Reardon (2012). "Animal models of atherosclerosis." Arterioscler Thromb Vasc Biol 32(5): 1104-1115.
- Ghesquiere, S. A., M. J. Gijbels, et al. (2005). "Macrophage-specific overexpression of group Ila sPLA2 increases atherosclerosis and enhances collagen deposition." J Lipid Res 46(2): 201-210.
- Giandomenico, V., M. Simonsson, et al. (2003). "Coactivator-dependent acetylation stabilizes members of the SREBP family of transcription factors." Mol Cell Biol 23(7): 2587-2599.
- Gimeno, D., G. L. Delclos, et al. (2011). "Association of CRP and IL-6 with lung function in a middle-aged population initially free from self-reported respiratory problems: the Whitehall II study." Eur J Epidemiol 26(2): 135-144.
- Glomset, J. A. (1968). "The plasma lecithins:cholesterol acyltransferase reaction." J Lipid Res 9(2): 155-167.
- Gora, S., G. Lambeau, et al. (2006). "The proinflammatory mediator Platelet Activating Factor is an effective substrate for human group X secreted phospholipase A2." Biochim Biophys Acta 1761(9): 1093-1099.
- Gora, S., C. Perret, et al. (2009). "Molecular and functional characterization of polymorphisms in the secreted phospholipase A2 group X gene: relevance to coronary artery disease." J Mol Med (Berl) 87(7): 723-733.
- Gordon, D. J., J. L. Probstfield, et al. (1989). "High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies." Circulation 79(1): 8-15.
- Gordon, T., W. B. Kannel, et al. (1974). "Death and coronary attacks in men after giving up cigarette smoking. A report from the Framingham study." Lancet 2(7893): 1345-1348.
- Goyal, T., S. Mitra, et al. (2012). "Current Concepts of the Role of Oxidized LDL Receptors in Atherosclerosis." Curr Atheroscler Rep.
- Graham, I. M. (2006). "The importance of total cardiovascular risk assessment in clinical practice." Eur J Gen Pract 12(4): 148-155.
- Granger, C. B., R. J. Goldberg, et al. (2003). "Predictors of hospital mortality in the global registry of acute coronary events." Arch Intern Med 163(19): 2345-2353.
- Gwilliam, B., V. Keeley, et al. (2011). "Development of prognosis in palliative care study (PiPS) predictor models to improve prognostication in advanced cancer: prospective cohort study." BMJ 343: d4920.
- Haffner, S. M., S. Lehto, et al. (1998). "Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction." N Engl J Med 339(4): 229-234.
- Haidich, A. B., C. Birtsou, et al. (2011). "The quality of safety reporting in trials is still suboptimal: survey of major general medical journals." J Clin Epidemiol 64(2): 124-135.

- Hallstrand, T. S., E. Y. Chi, et al. (2007). "Secreted phospholipase A2 group X overexpression in asthma and bronchial hyperresponsiveness." Am J Respir Crit Care Med 176(11): 1072-1078.
- Hanasaki, K., K. Yamada, et al. (2002). "Potent modification of low density lipoprotein by group X secretory phospholipase A2 is linked to macrophage foam cell formation." J Biol Chem 277(32): 29116-29124.
- Hansson, G. K. (2009). "Atherosclerosis--an immune disease: The Anitschkov Lecture 2007." Atherosclerosis 202(1): 2-10.
- Hansson, G. K. (2009). "Inflammatory mechanisms in atherosclerosis." J Thromb Haemost 7 Suppl 1: 328-331.
- Hansson, L., A. Zanchetti, et al. (1998). "Effects of intensive blood-pressure lowering and low-dose aspirin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomised trial. HOT Study Group." Lancet 351(9118): 1755-1762.
- Hase, M., M. Tanaka, et al. (2002). "Reduction in the extent of atherosclerosis in apolipoprotein E-deficient mice induced by electroporation-mediated transfer of the human plasma platelet-activating factor acetylhydrolase gene into skeletal muscle." Prostaglandins Other Lipid Mediat 70(1-2): 107-118.
- Hawe, E., P. J. Talmud, et al. (2003). "Family history is a coronary heart disease risk factor in the Second Northwick Park Heart Study." Ann Hum Genet 67(Pt 2): 97-106.
- Henderson, W. R., Jr., E. Y. Chi, et al. (2007). "Importance of group X-secreted phospholipase A2 in allergen-induced airway inflammation and remodeling in a mouse asthma model." J Exp Med 204(4): 865-877.
- Henderson, W. R., Jr., R. C. Oslund, et al. (2011). "Blockade of human group X secreted phospholipase A2 (GX-sPLA2)-induced airway inflammation and hyperresponsiveness in a mouse asthma model by a selective GX-sPLA2 inhibitor." J Biol Chem 286(32): 28049-28055.
- Henskens, L. H., N. Broos, et al. (2011). "A patient with bilateral shoulder and pelvic girdle aching." BMJ 343: d6233.
- Hermansson, M., K. Hokynar, et al. (2011). "Mechanisms of glycerophospholipid homeostasis in mammalian cells." Prog Lipid Res 50(3): 240-257.
- Higgins, J. P., S. G. Thompson, et al. (2003). "Measuring inconsistency in meta-analyses." BMJ 327(7414): 557-560.
- Hingorani, A. and S. Humphries (2005). "Nature's randomised trials." Lancet 366(9501): 1906-1908.
- Hjorth, S., H. J. Bengtsson, et al. (2000). "Serotonin autoreceptor function and antidepressant drug action." J Psychopharmacol 14(2): 177-185.
- Hofman, A., M. M. Breteler, et al. (2007). "The Rotterdam Study: objectives and design update." Eur J Epidemiol 22(11): 819-829.
- Holdt, L. M., F. Beutner, et al. (2010). "ANRIL expression is associated with atherosclerosis risk at chromosome 9p21." Arterioscler Thromb Vasc Biol 30(3): 620-627.
- Humphries, S. E., F. Drenos, et al. (2010). "Coronary heart disease risk prediction in the era of genome-wide association studies: current status and what the future holds." Circulation 121(20): 2235-2248.
- Humphries, S. E., L. A. Luong, et al. (2001). "The interleukin-6 -174 G/C promoter polymorphism is associated with risk of coronary heart disease and systolic blood pressure in healthy men." Eur Heart J 22(24): 2243-2252.
- Hurt-Camejo, E., S. Andersen, et al. (1997). "Localization of nonpancreatic secretory phospholipase A2 in normal and atherosclerotic arteries. Activity of the isolated enzyme on low-density lipoproteins." Arterioscler Thromb Vasc Biol 17(2): 300-309.
- Hurt-Camejo, E., G. Camejo, et al. (2001). "Phospholipase A(2) in vascular disease." Circ Res 89(4): 298-304.

- Hurt-Camejo, E., G. Camejo, et al. (2000). "Phospholipase A2 and small, dense low-density lipoprotein." *Curr Opin Lipidol* 11(5): 465-471.
- Innocenti, F., G. M. Cooper, et al. (2011). "Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue." *PLoS Genet* 7(5): e1002078.
- Interleukin-6 Receptor Mendelian Randomisation Analysis, C., A. D. Hingorani, et al. (2012). "The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis." *Lancet* 379(9822): 1214-1224.
- International Human Genome Sequencing, C. (2004). "Finishing the euchromatic sequence of the human genome." *Nature* 431(7011): 931-945.
- Investigators, G. (2001). "Rationale and design of the GRACE (Global Registry of Acute Coronary Events) Project: a multinational registry of patients hospitalized with acute coronary syndromes." *Am Heart J* 141(2): 190-199.
- Ishimoto, Y., K. Yamada, et al. (2003). "Group V and X secretory phospholipase A(2)s-induced modification of high-density lipoprotein linked to the reduction of its antiatherogenic functions." *Biochim Biophys Acta* 1642(3): 129-138.
- Ivandic, B., L. W. Castellani, et al. (1999). "Role of group II secretory phospholipase A2 in atherosclerosis: 1. Increased atherogenesis and altered lipoproteins in transgenic mice expressing group IIA phospholipase A2." *Arterioscler Thromb Vasc Biol* 19(5): 1284-1290.
- Jain, M. K. and P. M. Ridker (2005). "Anti-inflammatory effects of statins: clinical evidence and basic mechanisms." *Nat Rev Drug Discov* 4(12): 977-987.
- Jaross, W., R. Eckey, et al. (2002). "Biological effects of secretory phospholipase A(2) group IIA on lipoproteins and in atherogenesis." *Eur J Clin Invest* 32(6): 383-393.
- Jaulmes, A., B. Janvier, et al. (2005). "Autocrine and paracrine transcriptional regulation of type IIA secretory phospholipase A2 gene in vascular smooth muscle cells." *Arterioscler Thromb Vasc Biol* 25(6): 1161-1167.
- Jaulmes, A., S. Thierry, et al. (2006). "Activation of sPLA2-IIA and PGE2 production by high mobility group protein B1 in vascular smooth muscle cells sensitized by IL-1beta." *FASEB J* 20(10): 1727-1729.
- Jemel, I., H. Li, et al. (2011). "Group X secreted phospholipase A2 proenzyme is matured by a furin-like proprotein convertase and releases arachidonic acid inside of human HEK293 cells." *J Biol Chem* 286(42): 36509-36521.
- Jensen, G. B. and J. Hampton (2007). "Early termination of drug trials." *BMJ* 334(7589): 326.
- Jensen, M. D., W. Sheng, et al. (2009). "Involvement of oxidative pathways in cytokine-induced secretory phospholipase A2-IIA in astrocytes." *Neurochem Int* 55(6): 362-368.
- Jones, A., S. S. Dhamrait, et al. (2003). "Genetic variants of angiotensin II receptors and cardiovascular risk in hypertension." *Hypertension* 42(4): 500-506.
- Jonsson-Rylander, A. C., S. Lundin, et al. (2008). "Role of secretory phospholipases in atherogenesis." *Curr Atheroscler Rep* 10(3): 252-259.
- Jousilahti, P., E. Vartiainen, et al. (1999). "Sex, age, cardiovascular risk factors, and coronary heart disease: a prospective follow-up study of 14 786 middle-aged men and women in Finland." *Circulation* 99(9): 1165-1172.
- Jun, M., C. Foote, et al. (2010). "Effects of fibrates on cardiovascular outcomes: a systematic review and meta-analysis." *Lancet* 375(9729): 1875-1884.
- Kafonek, S. D. (1994). "Postmenopausal hormone replacement therapy and cardiovascular risk reduction. A review." *Drugs* 47 Suppl 2: 16-24.
- Kannel, W. B. and D. L. McGee (1979). "Diabetes and cardiovascular risk factors: the Framingham study." *Circulation* 59(1): 8-13.
- Karabina, S. A., I. Brocheriou, et al. (2006). "Atherogenic properties of LDL particles modified by human group X secreted phospholipase A2 on human endothelial cell function." *FASEB J* 20(14): 2547-2549.

- Karabina, S. A., S. Gora, et al. (2010). "Extracellular phospholipases in atherosclerosis." Biochimie 92(6): 594-600.
- Katan, M. B. (1986). "Apolipoprotein E isoforms, serum cholesterol, and cancer." Lancet 1(8479): 507-508.
- Kathiresan, S., Q. Yang, et al. (2006). "Common genetic variation in five thrombosis genes and relations to plasma hemostatic protein level and cardiovascular disease risk." Arterioscler Thromb Vasc Biol 26(6): 1405-1412.
- Kawakami, A., M. Osaka, et al. (2008). "Toll-like receptor 2 mediates apolipoprotein CIII-induced monocyte activation." Circ Res 103(12): 1402-1409.
- Keavney, B. (2011). "C reactive protein and the risk of cardiovascular disease." BMJ 342: d144.
- Keavney, B., J. Danesh, et al. (2006). "Fibrinogen and coronary heart disease: test of causality by 'Mendelian randomization'." Int J Epidemiol 35(4): 935-943.
- Ken-Dror, G., J. A. Cooper, et al. (2011). "Free protein S level as a risk factor for coronary heart disease and stroke in a prospective cohort study of healthy United Kingdom men." Am J Epidemiol 174(8): 958-968.
- Kennedy, B. P., P. Payette, et al. (1995). "A natural disruption of the secretory group II phospholipase A2 gene in inbred mouse strains." J Biol Chem 270(38): 22378-22385.
- Kim, J. O., B. K. Chakrabarti, et al. (2007). "Lysis of human immunodeficiency virus type 1 by a specific secreted human phospholipase A2." J Virol 81(3): 1444-1450.
- Kim, Y. J., K. P. Kim, et al. (2002). "Group V phospholipase A2 induces leukotriene biosynthesis in human neutrophils through the activation of group IVA phospholipase A2." J Biol Chem 277(39): 36479-36488.
- Kimmel, S. E., J. A. Berlin, et al. (2004). "The effects of nonselective non-aspirin non-steroidal anti-inflammatory medications on the risk of nonfatal myocardial infarction and their interaction with aspirin." J Am Coll Cardiol 43(6): 985-990.
- Kimura-Matsumoto, M., Y. Ishikawa, et al. (2008). "Expression of secretory phospholipase A2s in human atherosclerosis development." Atherosclerosis 196(1): 81-91.
- Kindel, T., D. M. Lee, et al. (2010). "The mechanism of the formation and secretion of chylomicrons." Atheroscler Suppl 11(1): 11-16.
- Kjekshus, J. and T. R. Pedersen (1995). "Reducing the risk of coronary events: evidence from the Scandinavian Simvastatin Survival Study (4S)." Am J Cardiol 76(9): 64C-68C.
- Kleber, M. E., R. L. Wolfert, et al. (2011). "Lipoprotein associated phospholipase A2 concentration predicts total and cardiovascular mortality independently of established risk factors (The Ludwigshafen Risk and Cardiovascular Health Study)." Clin Lab 57(9-10): 659-667.
- Koduri, R. S., J. O. Gronroos, et al. (2002). "Bactericidal properties of human and murine groups I, II, V, X, and XII secreted phospholipases A(2)." J Biol Chem 277(8): 5849-5857.
- Koenig, W. and N. Khuseyinova (2009). "Lipoprotein-associated and secretory phospholipase A2 in cardiovascular disease: the epidemiological evidence." Cardiovasc Drugs Ther 23(1): 85-92.
- Koenig, W., C. Y. Vossen, et al. (2009). "Association between type II secretory phospholipase A2 plasma concentrations and activity and cardiovascular events in patients with coronary heart disease." Eur Heart J 30(22): 2742-2748.
- Krieger, M. and J. Herz (1994). "Structures and functions of multiligand lipoprotein receptors: macrophage scavenger receptors and LDL receptor-related protein (LRP)." Annu Rev Biochem 63: 601-637.
- Kudo, I. and M. Murakami (2002). "Phospholipase A2 enzymes." Prostaglandins Other Lipid Mediat 68-69: 3-58.
- Kugiyama, K., Y. Ota, et al. (1999). "Circulating levels of secretory type II phospholipase A(2) predict coronary events in patients with coronary artery disease." Circulation 100(12): 1280-1284.
- Kuller, L. H. and E. N. Meilahn (1996). "Risk factors for cardiovascular disease among women." Curr Opin Lipidol 7(4): 203-208.

- Kutyavin, I. V., E. A. Lukhtanov, et al. (1997). "Oligonucleotides with conjugated dihydropyrroloindole tripeptides: base composition and backbone effects on hybridization." Nucleic Acids Res 25(18): 3718-3723.
- Kwiterovich, P. O., Jr. (2000). "The metabolic pathways of high-density lipoprotein, low-density lipoprotein, and triglycerides: a current review." Am J Cardiol 86(12A): 5L-10L.
- Lambeau, G. and M. H. Gelb (2008). "Biochemistry and physiology of mammalian secreted phospholipases A2." Annu Rev Biochem 77: 495-520.
- Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." Nature 409(6822): 860-921.
- Langer, H. F. and M. Gawaz (2008). "Platelet-vessel wall interactions in atherosclerotic disease." Thromb Haemost 99(3): 480-486.
- Lawlor, D. A., C. Bedford, et al. (2002). "Agreement between measured and self-reported weight in older women. Results from the British Women's Heart and Health Study." Age Ageing 31(3): 169-174.
- Lawlor, D. A., C. Bedford, et al. (2003). "Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study." J Epidemiol Community Health 57(2): 134-140.
- Lawlor, D. A., R. M. Harbord, et al. (2008). "Mendelian randomization: using genes as instruments for making causal inferences in epidemiology." Stat Med 27(8): 1133-1163.
- Leipold, H., C. Worda, et al. (2005). "Severe diabetic fetopathy despite strict metabolic control." Wien Klin Wochenschr 117(15-16): 561-564.
- Leite, J. O., U. Vaishnav, et al. (2009). "A-002 (Varespladib), a phospholipase A2 inhibitor, reduces atherosclerosis in guinea pigs." BMC Cardiovasc Disord 9: 7.
- Leonarduzzi, G., P. Gamba, et al. (2011). "Inflammation-related gene expression by lipid oxidation-derived products in the progression of atherosclerosis." Free Radic Biol Med.
- Lewington, S., R. Clarke, et al. (2002). "Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies." Lancet 360(9349): 1903-1913.
- Libby, P. (1995). "Molecular bases of the acute coronary syndromes." Circulation 91(11): 2844-2850.
- Libby, P., P. M. Ridker, et al. (2011). "Progress and challenges in translating the biology of atherosclerosis." Nature 473(7347): 317-325.
- Ligthelm, R. J., V. Borzi, et al. (2007). "Importance of observational studies in clinical practice." Clin Ther 29 Spec No: 1284-1292.
- Lin, Y., S. S. Mousa, et al. (2010). "Current status and future directions in lipid management: emphasizing low-density lipoproteins, high-density lipoproteins, and triglycerides as targets for therapy." Vasc Health Risk Manag 6: 73-85.
- Lind, L., T. Simon, et al. (2012). "Circulating levels of secretory- and lipoprotein-associated phospholipase A2 activities: relation to atherosclerotic plaques and future all-cause mortality." Eur Heart J.
- Liu, J. L., N. Maniadakis, et al. (2002). "The economic burden of coronary heart disease in the UK." Heart 88(6): 597-603.
- Lloyd-Jones, D. M., M. G. Larson, et al. (1999). "Lifetime risk of developing coronary heart disease." Lancet 353(9147): 89-92.
- Lloyd-Jones, D. M., E. P. Leip, et al. (2005). "Novel approach to examining first cardiovascular events after hypertension onset." Hypertension 45(1): 39-45.
- Lp, P. L. A. S. C., A. Thompson, et al. (2010). "Lipoprotein-associated phospholipase A(2) and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies." Lancet 375(9725): 1536-1544.
- Lu, J. T. and M. A. Creager (2004). "The relationship of cigarette smoking to peripheral arterial disease." Rev Cardiovasc Med 5(4): 189-193.

- Lu, M. and J. Y. Shyy (2006). "Sterol regulatory element-binding protein 1 is negatively modulated by PKA phosphorylation." Am J Physiol Cell Physiol 290(6): C1477-1486.
- Luscher, T. F. and M. Barton (1997). "Biology of the endothelium." Clin Cardiol 20(11 Suppl 2): II-3-10.
- Lusis, A. J. (2000). "Atherosclerosis." Nature 407(6801): 233-241.
- Lusis, A. J. (2012). "Genetics of atherosclerosis." Trends Genet 28(6): 267-275.
- Lusis, A. J., R. Mar, et al. (2004). "Genetics of atherosclerosis." Annu Rev Genomics Hum Genet 5: 189-218.
- Mahley, R. W., T. L. Innerarity, et al. (1984). "Plasma lipoproteins: apolipoprotein structure and function." J Lipid Res 25(12): 1277-1294.
- Majeed, F. and M. Miller (2008). "Low high-density lipoprotein cholesterol: an important consideration in coronary heart disease risk assessment." Curr Opin Endocrinol Diabetes Obes 15(2): 175-181.
- Mallat, Z., J. Benessiano, et al. (2007). "Circulating secretory phospholipase A2 activity and risk of incident coronary events in healthy men and women: the EPIC-Norfolk study." Arterioscler Thromb Vasc Biol 27(5): 1177-1183.
- Mallat, Z., G. Lambeau, et al. (2010). "Lipoprotein-associated and secreted phospholipases A in cardiovascular disease: roles as biological effectors and biomarkers." Circulation 122(21): 2183-2200.
- Mallat, Z., P. G. Steg, et al. (2005). "Circulating secretory phospholipase A2 activity predicts recurrent events in patients with severe acute coronary syndromes." J Am Coll Cardiol 46(7): 1249-1257.
- Mann, C. J. (2003). "Observational research methods. Research design II: cohort, cross sectional, and case-control studies." Emerg Med J 20(1): 54-60.
- Manolio, T. A. (2010). "Genomewide association studies and assessment of the risk of disease." N Engl J Med 363(2): 166-176.
- Maraldi, N. M., G. Mazzotti, et al. (1992). "Morphological evidence of function-related localization of phospholipids in the cell nucleus." Adv Enzyme Regul 32: 73-90.
- Marchini, J., B. Howie, et al. (2007). "A new multipoint method for genome-wide association studies by imputation of genotypes." Nat Genet 39(7): 906-913.
- Marenberg, M. E., N. Risch, et al. (1994). "Genetic susceptibility to death from coronary heart disease in a study of twins." N Engl J Med 330(15): 1041-1046.
- Marian, A. J. (2010). "The personal genome and the practice of cardiovascular medicine." Methodist Debaque Cardiovasc J 6(4): 13-20.
- Marmot, M. and E. Brunner (2005). "Cohort Profile: the Whitehall II study." Int J Epidemiol 34(2): 251-256.
- Marquart, T. J., R. M. Allen, et al. (2010). "miR-33 links SREBP-2 induction to repression of sterol transporters." Proc Natl Acad Sci U S A 107(27): 12228-12232.
- Martin, S. S., R. S. Blumenthal, et al. (2012). "LDL cholesterol: the lower the better." Med Clin North Am 96(1): 13-26.
- Mason, R. P. (2011). "Optimal therapeutic strategy for treating patients with hypertension and atherosclerosis: focus on olmesartan medoxomil." Vasc Health Risk Manag 7: 405-416.
- Masuda, S., M. Murakami, et al. (2005). "Diverse cellular localizations of secretory phospholipase A2 enzymes in several human tissues." Biochim Biophys Acta 1736(3): 200-210.
- Masuda, S., M. Murakami, et al. (2005). "Various secretory phospholipase A2 enzymes are expressed in rheumatoid arthritis and augment prostaglandin production in cultured synovial cells." FEBS J 272(3): 655-672.
- Masuda, S., M. Murakami, et al. (2005). "Expression of secretory phospholipase A2 enzymes in lungs of humans with pneumonia and their potential prostaglandin-synthetic function in human lung-derived cells." Biochem J 387(Pt 1): 27-38.

- Mathers, C. D. and D. Loncar (2006). "Projections of global mortality and burden of disease from 2002 to 2030." *PLoS Med* 3(11): e442.
- Matthews, K. A., S. A. Shumaker, et al. (1997). "Women's health initiative. Why now? What is it? What's new?" *Am Psychol* 52(2): 101-116.
- McDermott-Roe, C., J. Ye, et al. (2011). "Endonuclease G is a novel determinant of cardiac hypertrophy and mitochondrial function." *Nature* 478(7367): 114-118.
- Meade, T. W., H. C. Wilkes, et al. (1988). "Randomized controlled trial of low dose warfarin in the primary prevention of ischaemic heart disease in men at high risk: design and pilot study." *Eur Heart J* 9(8): 836-843.
- Menschikowski, M., A. Hagelgans, et al. (2008). "Differential expression of secretory phospholipases A2 in normal and malignant prostate cell lines: regulation by cytokines, cell signaling pathways, and epigenetic mechanisms." *Neoplasia* 10(3): 279-286.
- Menschikowski, M., A. Hagelgans, et al. (2006). "Secretory phospholipase A2 of group IIA: is it an offensive or a defensive player during atherosclerosis and other inflammatory diseases?" *Prostaglandins Other Lipid Mediat* 79(1-2): 1-33.
- Mihelich, E. D. and R. W. Schevitz (1999). "Structure-based design of a new class of anti-inflammatory drugs: secretory phospholipase A(2) inhibitors, SPI." *Biochim Biophys Acta* 1441(2-3): 223-228.
- Miller, G. J., K. A. Bauer, et al. (1996). "Increased activation of the haemostatic system in men at high risk of fatal coronary heart disease." *Thromb Haemost* 75(5): 767-771.
- Minelli, C., J. R. Thompson, et al. (2004). "An integrated approach to the meta-analysis of genetic association studies using Mendelian randomization." *Am J Epidemiol* 160(5): 445-452.
- Mitsuishi, M., S. Masuda, et al. (2006). "Group V and X secretory phospholipase A2 prevents adenoviral infection in mammalian cells." *Biochem J* 393(Pt 1): 97-106.
- Moher, D., A. Liberati, et al. (2009). "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement." *PLoS Med* 6(7): e1000097.
- Moon, R. T., J. D. Brown, et al. (1997). "Structurally related receptors and antagonists compete for secreted Wnt ligands." *Cell* 88(6): 725-728.
- Moore, H. C., M. Johnston, et al. (2011). "An evolutionarily conserved, alternatively spliced, intron in the p68/DDX5 DEAD-box RNA helicase gene encodes a novel miRNA." *RNA* 17(4): 555-562.
- Moreno, P. R., J. Sanz, et al. (2009). "Promoting mechanisms of vascular health: circulating progenitor cells, angiogenesis, and reverse cholesterol transport." *J Am Coll Cardiol* 53(25): 2315-2323.
- Morrow, D. A., B. M. Scirica, et al. (2007). "Effects of ranolazine on recurrent cardiovascular events in patients with non-ST-elevation acute coronary syndromes: the MERLIN-TIMI 36 randomized trial." *JAMA* 297(16): 1775-1783.
- Moses, G. S., M. D. Jensen, et al. (2006). "Secretory PLA2-IIA: a new inflammatory factor for Alzheimer's disease." *J Neuroinflammation* 3: 28.
- Moss, S. E., R. Klein, et al. (1991). "Cause-specific mortality in a population-based study of diabetes." *Am J Public Health* 81(9): 1158-1162.
- Mudau, M., A. Genis, et al. (2012). "Endothelial dysfunction: the early predictor of atherosclerosis." *Cardiovasc J Afr* 23(4): 222-231.
- Mullis, K., F. Faloona, et al. (1986). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction." *Cold Spring Harb Symp Quant Biol* 51 Pt 1: 263-273.
- Munoz, N. M., A. Y. Meliton, et al. (2007). "Deletion of secretory group V phospholipase A2 attenuates cell migration and airway hyperresponsiveness in immunosensitized mice." *J Immunol* 179(7): 4800-4807.
- Munoz, N. M., A. Y. Meliton, et al. (2009). "Secretory group V phospholipase A2 regulates acute lung injury and neutrophilic inflammation caused by LPS in mice." *Am J Physiol Lung Cell Mol Physiol* 296(6): L879-887.

- Murakami, M., T. Kambe, et al. (1999). "Functional coupling between various phospholipase A2s and cyclooxygenases in immediate and delayed prostanoid biosynthetic pathways." J Biol Chem 274(5): 3103-3115.
- Murakami, M. and I. Kudo (1999). "[Functional coupling between phospholipase A2 and cyclooxygenase]." Tanpakushitsu Kakusan Koso 44(8 Suppl): 1146-1156.
- Murakami, M. and I. Kudo (2003). "New phospholipase A(2) isozymes with a potential role in atherosclerosis." Curr Opin Lipidol 14(5): 431-436.
- Murakami, M., I. Kudo, et al. (1989). "In vivo release and clearance of rat platelet phospholipase A2." Biochim Biophys Acta 1005(3): 270-276.
- Murakami, M., Y. Taketomi, et al. (2011). "Secreted phospholipase A2 revisited." J Biochem 150(3): 233-255.
- Murakami, Y., S. Matsufuji, et al. (1989). "Properties and fluctuations in vivo of rat liver antizyme inhibitor." Biochem J 259(3): 839-845.
- Naito, Y., J. Fukata, et al. (1989). "Adrenocorticotrophic hormone-releasing activities of interleukins in a homologous in vivo system." Biochem Biophys Res Commun 164(3): 1262-1267.
- Nakos, G., E. Kitsioulis, et al. (2005). "Phospholipases A2 and platelet-activating-factor acetylhydrolase in patients with acute respiratory distress syndrome." Crit Care Med 33(4): 772-779.
- Nicholls, S. J., M. A. Cavender, et al. (2011). "Inhibition of Secretory Phospholipase A(2) in Patients with Acute Coronary Syndromes: Rationale and Design of the Vascular Inflammation Suppression to Treat Acute Coronary Syndrome for 16 Weeks (VISTA-16) Trial." Cardiovasc Drugs Ther.
- Nichols, A. V. (1969). "Functions and interrelationships of different classes of plasma lipoproteins." Proc Natl Acad Sci U S A 64(3): 1128-1137.
- Nordestgaard, B. G., B. Agerholm-Larsen, et al. (1996). "A prospective cardiovascular population study used in genetic epidemiology. The Copenhagen City Heart Study." Scand J Clin Lab Invest Suppl 226: 65-71.
- Ockene, I. S. and N. H. Miller (1997). "Cigarette smoking, cardiovascular disease, and stroke: a statement for healthcare professionals from the American Heart Association. American Heart Association Task Force on Risk Reduction." Circulation 96(9): 3243-3247.
- Ockene, J. K., L. H. Kuller, et al. (1990). "The relationship of smoking cessation to coronary heart disease and lung cancer in the Multiple Risk Factor Intervention Trial (MRFIT)." Am J Public Health 80(8): 954-958.
- Oestvang, J., M. W. Anthonsen, et al. (2003). "Role of secretory and cytosolic phospholipase A(2) enzymes in lysophosphatidylcholine-stimulated monocyte arachidonic acid release." FEBS Lett 555(2): 257-262.
- Ohtsuki, M., Y. Taketomi, et al. (2006). "Transgenic expression of group V, but not group X, secreted phospholipase A2 in mice leads to neonatal lethality because of lung dysfunction." J Biol Chem 281(47): 36420-36433.
- Olive, K. P., M. A. Jacobetz, et al. (2009). "Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer." Science 324(5933): 1457-1461.
- Olivier, J. L., Q. Fan, et al. (1994). "Positive and negative hepatic regulation of the human type II phospholipase A2 gene." Biochemistry 33(23): 7134-7145.
- Oorni, K. and P. T. Kovanen (2009). "Lipoprotein modification by secretory phospholipase A(2) enzymes contributes to the initiation and progression of atherosclerosis." Curr Opin Lipidol 20(5): 421-427.
- Orso, E., M. Grandl, et al. (2011). "Oxidized LDL-induced endolysosomal phospholipidosis and enzymatically modified LDL-induced foam cell formation determine specific lipid species modulation in human macrophages." Chem Phys Lipids 164(6): 479-487.
- Packard, C., M. Caslake, et al. (2000). "The role of small, dense low density lipoprotein (LDL): a new look." Int J Cardiol 74 Suppl 1: S17-22.

- Panayiotou, A., A. Nicolaides, et al. (2009). "Serum total homocysteine, folate, 5,10-methylenetetrahydrofolate reductase (MTHFR) 677C-->T genotype and subclinical atherosclerosis." Expert Opin Ther Targets 13(1): 1-11.
- Paradon, M., C. Salvat, et al. (1998). "An SP1-like 5'-GACCACGCC-3' sequence is critical for activity of the inflammatory phospholipase A2 promoter and binds several non-zinc finger proteins." Eur J Biochem 258(1): 113-122.
- Park, Y. M., R. K. S, et al. (2012). "Insulin promotes macrophage foam cell formation: potential implications in diabetes-related atherosclerosis." Lab Invest 92(8): 1171-1180.
- Parra, A., E. Perez-Pasten, et al. (1974). "[Hyperinsulinemia in children with chemical diabetes mellitus]." Arch Invest Med (Mex) 5(2): 117-124.
- Patsch, J. R., G. Miesenbock, et al. (1992). "Relation of triglyceride metabolism and coronary artery disease. Studies in the postprandial state." Arterioscler Thromb 12(11): 1336-1345.
- Peilot, H., B. Rosengren, et al. (2000). "Interferon-gamma induces secretory group IIA phospholipase A2 in human arterial smooth muscle cells. Involvement of cell differentiation, STAT-3 activation, and modulation by other cytokines." J Biol Chem 275(30): 22895-22904.
- Pekkanen, J., S. Linn, et al. (1990). "Ten-year mortality from cardiovascular disease in relation to cholesterol level among men with and without preexisting cardiovascular disease." N Engl J Med 322(24): 1700-1707.
- Pentikainen, M. O., K. Oorni, et al. (2000). "Modified LDL - trigger of atherosclerosis and inflammation in the arterial intima." J Intern Med 247(3): 359-370.
- Petretto, E., J. Mangion, et al. (2006). "Heritability and tissue specificity of expression quantitative trait loci." PLoS Genet 2(10): e172.
- Pitha, J., J. A. Hubáček, et al. (2007). "Genetic determination of the prognosis in survivors of acute coronary syndromes. Study design and rationale for a multicenter study." Cor et Vasa 49(4): 134-137.
- Pohjola-Sintonen, S., A. Rissanen, et al. (1998). "Family history as a risk factor of coronary heart disease in patients under 60 years of age." Eur Heart J 19(2): 235-239.
- Ponugoti, B., D. H. Kim, et al. (2010). "SIRT1 deacetylates and inhibits SREBP-1C activity in regulation of hepatic lipid metabolism." J Biol Chem 285(44): 33959-33970.
- Prentice, R. L., S. Paczesny, et al. (2010). "Novel proteins associated with risk for coronary heart disease or stroke among postmenopausal women identified by in-depth plasma proteome profiling." Genome Med 2(7): 48.
- Pruzanski, W. (1985). "Viscosity and hyperviscosity: application of hemorheological concepts to rheumatology." J Rheumatol 12(3): 401-402.
- Pruzanski, W., E. C. Keystone, et al. (1988). "Serum phospholipase A2 correlates with disease activity in rheumatoid arthritis." J Rheumatol 15(9): 1351-1355.
- Pruzanski, W., G. Lambeau, et al. (2007). "Hydrolysis of minor glycerophospholipids of plasma lipoproteins by human group IIA, V and X secretory phospholipases A2." Biochim Biophys Acta 1771(1): 5-19.
- Pruzanski, W., L. Lambeau, et al. (2005). "Differential hydrolysis of molecular species of lipoprotein phosphatidylcholine by groups IIA, V and X secretory phospholipases A2." Biochim Biophys Acta 1736(1): 38-50.
- Pruzanski, W. and P. Vadas (1991). "Phospholipase A2--a mediator between proximal and distal effectors of inflammation." Immunol Today 12(5): 143-146.
- Pruzanski, W., P. Vadas, et al. (1985). "Phospholipase A2 activity in sera and synovial fluids in rheumatoid arthritis and osteoarthritis. Its possible role as a proinflammatory enzyme." J Rheumatol 12(2): 211-216.
- Quint, J. K., G. C. Donaldson, et al. (2011). "SERPINA1 11478G-->A variant, serum alpha1-antitrypsin, exacerbation frequency and FEV1 decline in COPD." Thorax 66(5): 418-424.

- Rader, D. J. and A. Daugherty (2008). "Translating molecular discoveries into new therapies for atherosclerosis." *Nature* 451(7181): 904-913.
- Ray, W. A., C. M. Stein, et al. (2002). "Non-steroidal anti-inflammatory drugs and risk of serious coronary heart disease: an observational cohort study." *Lancet* 359(9301): 118-123.
- Reilly, M. P., M. Li, et al. (2011). "Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies." *Lancet* 377(9763): 383-392.
- Reinhart, W. H. (2003). "Fibrinogen--marker or mediator of vascular disease?" *Vasc Med* 8(3): 211-216.
- Reynolds, L. J., L. L. Hughes, et al. (1994). "1-Hexadecyl-2-arachidonoylthio-2-deoxy-sn-glycero-3-phosphorylcholine as a substrate for the microtiterplate assay of human cytosolic phospholipase A2." *Anal Biochem* 217(1): 25-32.
- Rosengren, A., M. Hagman, et al. (1997). "Serum cholesterol and long-term prognosis in middle-aged men with myocardial infarction and angina pectoris. A 16-year follow-up of the Primary Prevention Study in Goteborg, Sweden." *Eur Heart J* 18(5): 754-761.
- Rosengren, B., A. C. Jonsson-Rylander, et al. (2006). "Distinctiveness of secretory phospholipase A2 group IIA and V suggesting unique roles in atherosclerosis." *Biochim Biophys Acta* 1761(11): 1301-1308.
- Rosengren, B., H. Peilot, et al. (2006). "Secretory phospholipase A2 group V: lesion distribution, activation by arterial proteoglycans, and induction in aorta by a Western diet." *Arterioscler Thromb Vasc Biol* 26(7): 1579-1585.
- Rosenson, R. S., M. Elliott, et al. (2011). "Randomized trial of an inhibitor of secretory phospholipase A2 on atherogenic lipoprotein subclasses in statin-treated patients with coronary heart disease." *Eur Heart J* 32(8): 999-1005.
- Rosenson, R. S., H. Fraser, et al. (2010). "Varespladib methyl in cardiovascular disease." *Expert Opin Investig Drugs* 19(10): 1245-1255.
- Rosenson, R. S. and M. H. Gelb (2009). "Secretory phospholipase A2: a multifaceted family of proatherogenic enzymes." *Curr Cardiol Rep* 11(6): 445-451.
- Rosenson, R. S., C. Hislop, et al. (2010). "Effects of varespladib methyl on biomarkers and major cardiovascular events in acute coronary syndrome patients." *J Am Coll Cardiol* 56(14): 1079-1088.
- Rosenson, R. S., C. Hislop, et al. (2009). "Effects of 1-H-indole-3-glyoxamide (A-002) on concentration of secretory phospholipase A2 (PLASMA study): a phase II double-blind, randomised, placebo-controlled trial." *Lancet* 373(9664): 649-658.
- Rosenson, R. S. and E. Hurt-Camejo (2012). "Phospholipase A2 enzymes and the risk of atherosclerosis." *Eur Heart J*.
- Ross, R. (1999). "Atherosclerosis--an inflammatory disease." *N Engl J Med* 340(2): 115-126.
- Rothenbacher, D., W. Koenig, et al. (2006). "Comparison of N-terminal pro-B-natriuretic peptide, C-reactive protein, and creatinine clearance for prognosis in patients with known coronary heart disease." *Arch Intern Med* 166(22): 2455-2460.
- Rye, K. A. and P. J. Barter (2012). "Predictive value of different HDL particles for the protection against or risk of coronary heart disease." *Biochim Biophys Acta* 1821(3): 473-480.
- Sachidanandam, R., D. Weissman, et al. (2001). "A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms." *Nature* 409(6822): 928-933.
- Saiga, A., Y. Morioka, et al. (2001). "Group X secretory phospholipase A(2) induces potent productions of various lipid mediators in mouse peritoneal macrophages." *Biochim Biophys Acta* 1530(1): 67-76.
- Saksela, O. and D. B. Rifkin (1988). "Cell-associated plasminogen activation: regulation and physiological functions." *Annu Rev Cell Biol* 4: 93-126.

- Samani, N. J., P. Burton, et al. (2005). "A genomewide linkage study of 1,933 families affected by premature coronary artery disease: The British Heart Foundation (BHF) Family Heart Study." Am J Hum Genet 77(6): 1011-1020.
- Sampietro, M. L., D. Pons, et al. (2009). "A genome wide association analysis in the GENDER study." Neth Heart J 17(6): 262-264.
- Sartipy, P., B. Johansen, et al. (2000). "Molecular basis for the association of group IIA phospholipase A(2) and decorin in human atherosclerotic lesions." Circ Res 86(6): 707-714.
- Sarwar, N., J. Danesh, et al. (2007). "Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies." Circulation 115(4): 450-458.
- Satake, Y., B. L. Diaz, et al. (2004). "Role of group V phospholipase A2 in zymosan-induced eicosanoid generation and vascular permeability revealed by targeted gene disruption." J Biol Chem 279(16): 16488-16494.
- Sato, H., Y. Isogai, et al. (2011). "Physiological roles of group X-secreted phospholipase A2 in reproduction, gastrointestinal phospholipid digestion, and neuronal function." J Biol Chem 286(13): 11632-11648.
- Sato, H., R. Kato, et al. (2008). "Analyses of group III secreted phospholipase A2 transgenic mice reveal potential participation of this enzyme in plasma lipoprotein modification, macrophage foam cell formation, and atherosclerosis." J Biol Chem 283(48): 33483-33497.
- Sawada, H., M. Murakami, et al. (1999). "Regulation of type V phospholipase A2 expression and function by proinflammatory stimuli." Eur J Biochem 263(3): 826-835.
- Schevitz, R. W., N. J. Bach, et al. (1995). "Structure-based design of the first potent and selective inhibitor of human non-pancreatic secretory phospholipase A2." Nat Struct Biol 2(6): 458-465.
- Seilhamer, J. J., W. Pruzanski, et al. (1989). "Cloning and recombinant expression of phospholipase A2 present in rheumatoid arthritic synovial fluid." J Biol Chem 264(10): 5335-5338.
- Shaper, A. G., S. J. Pocock, et al. (1981). "British Regional Heart Study: cardiovascular risk factors in middle-aged men in 24 towns." Br Med J (Clin Res Ed) 283(6285): 179-186.
- Shaposhnik, Z., X. Wang, et al. (2009). "The synergistic inhibition of atherogenesis in apoE^{-/-} mice between pravastatin and the sPLA2 inhibitor varespladib (A-002)." J Lipid Res 50(4): 623-629.
- Sheehan, N. A., V. Didelez, et al. (2008). "Mendelian randomisation and causal inference in observational epidemiology." PLoS Med 5(8): e177.
- Shen, B. W., A. M. Scanu, et al. (1977). "Structure of human serum lipoproteins inferred from compositional analysis." Proc Natl Acad Sci U S A 74(3): 837-841.
- Shepherd, J., G. J. Blauw, et al. (2002). "Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial." Lancet 360(9346): 1623-1630.
- Shih, D. M., Y. R. Xia, et al. (2000). "Combined serum paraoxonase knockout/apolipoprotein E knockout mice exhibit increased lipoprotein oxidation and atherosclerosis." J Biol Chem 275(23): 17527-17535.
- Simons, P. C., A. Algra, et al. (1999). "Second manifestations of ARterial disease (SMART) study: rationale and design." Eur J Epidemiol 15(9): 773-781.
- Singh, D. K., L. R. Gesquiere, et al. (2007). "Role of sphingomyelin and ceramide in the regulation of the activity and fatty acid specificity of group V secretory phospholipase A2." Arch Biochem Biophys 459(2): 280-287.
- Singh, D. K. and P. V. Subbaiah (2007). "Modulation of the activity and arachidonic acid selectivity of group X secretory phospholipase A2 by sphingolipids." J Lipid Res 48(3): 683-692.
- Singh, P., M. Singh, et al. (2007). "The ApoAI-CIII-AIV gene cluster and its relation to lipid levels in type 2 diabetes mellitus and coronary heart disease: determination of a novel susceptible haplotype." Diab Vasc Dis Res 4(2): 124-129.

- Singh, P., M. Singh, et al. (2008). "A novel haplotype in ApoA1-CIII-AIV gene region is detrimental to Northwest Indians with coronary heart disease." Int J Cardiol 130(3): e93-95.
- Sitia, S., L. Tomasoni, et al. (2010). "From endothelial dysfunction to atherosclerosis." Autoimmun Rev 9(12): 830-834.
- Six, D. A. and E. A. Dennis (2000). "The expanding superfamily of phospholipase A(2) enzymes: classification and characterization." Biochim Biophys Acta 1488(1-2): 1-19.
- Smart, B. P., Y. H. Pan, et al. (2004). "Inhibition of the complete set of mammalian secreted phospholipases A(2) by indole analogues: a structure-guided study." Bioorg Med Chem 12(7): 1737-1749.
- Smith, C. L., S. Anthony, et al. (2005). "Effects of ADMA upon gene expression: an insight into the pathophysiological significance of raised plasma ADMA." PLoS Med 2(10): e264.
- Smith, G. D., R. Harbord, et al. (2005). "Does elevated plasma fibrinogen increase the risk of coronary heart disease? Evidence from a meta-analysis of genetic association studies." Arterioscler Thromb Vasc Biol 25(10): 2228-2233.
- Snyder, D. W., N. J. Bach, et al. (1999). "Pharmacology of LY315920/S-5920, [[3-(aminooxoacetyl)-2-ethyl-1- (phenylmethyl)-1H-indol-4-yl]oxy] acetate, a potent and selective secretory phospholipase A2 inhibitor: A new class of anti-inflammatory drugs, SPI." J Pharmacol Exp Ther 288(3): 1117-1124.
- Sofat, R., A. D. Hingorani, et al. (2010). "Separating the mechanism-based and off-target actions of cholesteryl ester transfer protein inhibitors with CETP gene polymorphisms." Circulation 121(1): 52-62.
- Steinberg, D. (1987). "Lipoproteins and the pathogenesis of atherosclerosis." Circulation 76(3): 508-514.
- Suh, Y. and J. Vijg (2005). "SNP discovery in associating genetic variation with human disease phenotypes." Mutat Res 573(1-2): 41-53.
- Tabor, H. K., N. J. Risch, et al. (2002). "Candidate-gene approaches for studying complex genetic traits: practical considerations." Nat Rev Genet 3(5): 391-397.
- Takahashi, N., H. Murakami, et al. (2000). "Association of a polymorphism at the 5'-region of the angiotensin II type 1 receptor with hypertension." Ann Hum Genet 64(Pt 3): 197-205.
- Takayama, K., S. Hara, et al. (1991). "Detection of 14-kDa group II phospholipase A2 in human seminal plasma." Biochem Biophys Res Commun 178(3): 1505-1511.
- Tegos, T. J., E. Kalodiki, et al. (2001). "The genesis of atherosclerosis and risk factors: a review." Angiology 52(2): 89-98.
- Thayer, S. P., M. P. di Magliano, et al. (2003). "Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis." Nature 425(6960): 851-856.
- Tietge, U. J., D. Pratico, et al. (2005). "Macrophage-specific expression of group IIA sPLA2 results in accelerated atherogenesis by increasing oxidative stress." J Lipid Res 46(8): 1604-1614.
- Tischfield, J. A., Y. R. Xia, et al. (1996). "Low-molecular-weight, calcium-dependent phospholipase A2 genes are linked and map to homologous chromosome regions in mouse and human." Genomics 32(3): 328-333.
- Tobin, M. D., S. M. Raleigh, et al. (2005). "Association of WNK1 gene polymorphisms and haplotypes with ambulatory blood pressure in the general population." Circulation 112(22): 3423-3429.
- Touqui, L. and M. Alaoui-El-Azher (2001). "Mammalian secreted phospholipases A2 and their pathophysiological significance in inflammatory diseases." Curr Mol Med 1(6): 739-754.
- Tribler, L., L. T. Jensen, et al. (2007). "Increased expression and activity of group IIA and X secretory phospholipase A2 in peritumoral versus central colon carcinoma tissue." Anticancer Res 27(5A): 3179-3185.
- Triggiani, M., C. Calabrese, et al. (2000). "Metabolism of lipid mediators in human eosinophils." Chem Immunol 76: 77-98.

- Triggiani, M., F. Granata, et al. (2000). "Secretory phospholipases A2 induce beta-glucuronidase release and IL-6 production from human lung macrophages." J Immunol 164(9): 4908-4915.
- Triglyceride Coronary Disease Genetics, C., C. Emerging Risk Factors, et al. (2010). "Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies." Lancet 375(9726): 1634-1639.
- Trip, M. D., Y. M. Smulders, et al. (2002). "Frequent mutation in the ABCG6 gene (R1141X) is associated with a strong increase in the prevalence of coronary artery disease." Circulation 106(7): 773-775.
- Trousseau, A., S. Bernard, et al. (2009). "25-hydroxycholesterol provokes oligodendrocyte cell line apoptosis and stimulates the secreted phospholipase A2 type IIA via LXR beta and PXR." J Neurochem 109(4): 945-958.
- Tsai, C., V. Smider, et al. (2012). "Electrophoretic mobility shift assays for protein-DNA complexes involved in DNA repair." Methods Mol Biol 920: 53-78.
- Tsai, N. P., Y. L. Lin, et al. (2009). "MicroRNA mir-346 targets the 5'-untranslated region of receptor-interacting protein 140 (RIP140) mRNA and up-regulates its protein expression." Biochem J 424(3): 411-418.
- Tzoulaki, I., K. C. Siontis, et al. (2011). "Prognostic effect size of cardiovascular biomarkers in datasets from observational studies versus randomised trials: meta-epidemiology study." BMJ 343: d6829.
- Valensi, P., M. Benroubi, et al. (2008). "The IMPROVE study--a multinational, observational study in type 2 diabetes: baseline characteristics from eight national cohorts." Int J Clin Pract 62(11): 1809-1819.
- Van den Bruel, A., M. J. Thompson, et al. (2011). "Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review." BMJ 342: d3082.
- van Meer, G., D. R. Voelker, et al. (2008). "Membrane lipids: where they are and how they behave." Nat Rev Mol Cell Biol 9(2): 112-124.
- van Wieren-de Wijer, D. B., A. H. Maitland-van der Zee, et al. (2009). "Recruitment of participants through community pharmacies for a pharmacogenetic study of antihypertensive drug treatment." Pharm World Sci 31(2): 158-164.
- Veeranna, V., S. K. Zalawadiya, et al. (2011). "Homocysteine and reclassification of cardiovascular disease risk." J Am Coll Cardiol 58(10): 1025-1033.
- Vervoordeldonk, M. J., C. G. Schalkwijk, et al. (1996). "Effects of dexamethasone and transforming growth factor-beta 2 on group II phospholipase A2 mRNA and activity levels in interleukin 1 beta- and forskolin-stimulated mesangial cells." Biochem J 315 (Pt 2): 435-441.
- Voight, B. F., G. M. Peloso, et al. (2012). "Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study." Lancet.
- Wasling, C. (1999). "Role of the cardioprotective diet in preventing coronary heart disease." Br J Nurs 8(18): 1239-1248.
- Waydhas, C., D. Nast-Kolb, et al. (1989). "Prognostic value of serum phospholipase A in the multitraumatized patient." Klin Wochenschr 67(3): 203-206.
- Webb, N. R., M. A. Bostrom, et al. (2003). "Macrophage-expressed group IIA secretory phospholipase A2 increases atherosclerotic lesion formation in LDL receptor-deficient mice." Arterioscler Thromb Vasc Biol 23(2): 263-268.
- Weber, C. and H. Noels (2011). "Atherosclerosis: current pathogenesis and therapeutic options." Nat Med 17(11): 1410-1422.
- Wellcome Trust Case Control, C., C. Australo-Anglo-American Spondylitis, et al. (2007). "Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants." Nat Genet 39(11): 1329-1337.
- Wery, J. P., R. W. Schevitz, et al. (1991). "Structure of recombinant human rheumatoid arthritic synovial fluid phospholipase A2 at 2.2 Å resolution." Nature 352(6330): 79-82.

- Willer, C. J., S. Sanna, et al. (2008). "Newly identified loci that influence lipid concentrations and risk of coronary artery disease." Nat Genet 40(2): 161-169.
- Willerson, J. T. and P. M. Ridker (2004). "Inflammation as a cardiovascular risk factor." Circulation 109(21 Suppl 1): II2-10.
- Wootton, P. T., N. L. Arora, et al. (2007). "Tagging SNP haplotype analysis of the secretory PLA2-V gene, PLA2G5, shows strong association with LDL and oxLDL levels, suggesting functional distinction from sPLA2-IIA: results from the UDACS study." Hum Mol Genet 16(12): 1437-1444.
- Wootton, P. T., F. Drenos, et al. (2006). "Tagging-SNP haplotype analysis of the secretory PLA2IIa gene PLA2G2A shows strong association with serum levels of sPLA2IIa: results from the UDACS study." Hum Mol Genet 15(2): 355-361.
- Yang, X. (2012). "Use of functional genomics to identify candidate genes underlying human genetic association studies of vascular diseases." Arterioscler Thromb Vasc Biol 32(2): 216-222.
- Yu, K. C. and A. D. Cooper (2001). "Postprandial lipoproteins and atherosclerosis." Front Biosci 6: D332-354.
- Yusuf, S., S. Reddy, et al. (2001). "Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization." Circulation 104(22): 2746-2753.
- Zebrack, J. S., J. L. Anderson, et al. (2002). "Usefulness of high-sensitivity C-reactive protein in predicting long-term risk of death or acute myocardial infarction in patients with unstable or stable angina pectoris or acute myocardial infarction." Am J Cardiol 89(2): 145-149.
- Zhang, C. (2008). "The role of inflammatory cytokines in endothelial dysfunction." Basic Res Cardiol 103(5): 398-406.
- Zheng, C., C. Khoo, et al. (2008). "Dietary monounsaturated fat activates metabolic pathways for triglyceride-rich lipoproteins that involve apolipoproteins E and C-III." Am J Clin Nutr 88(2): 272-281.
- Zhou, W., S. K. Watt, et al. (2012). "Quantitative Attachment and Detachment of Bacterial Spores from Fine Wires through Continuous and Pulsed DC Electrophoretic Deposition." J Phys Chem B.

Web and Additional References

- F.B. Hu, [International Encyclopaedia of Public Health](#); pages 181-190
- GE Life Sciences: <http://www.gelifesciences.com/>
- GenWay Biotech Inc, http://www.genwaybio.com/product_info.php?products_id=45346
- PRISMA: <http://www.prisma-statement.org/statement.htm>
- Promega: <http://www.promega.com/>
- Qiagen: www.qiagen.com/
- Sigma-Aldrich: <http://www.sigmaaldrich.com/>

WHO Media Centre: <http://www.who.int/mediacentre/>

WHO: <http://www.who.int/mediacentre/>

Wise Owl: www.wiseowl.id.au

US Dept. of Health and Human Services. Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General. US Department of Health and Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 1989.

DHHS Publication (CDC) 89-8411 AND US Dept. of Health and Human Services. The Health Benefits of Smoking Cessation. A Report of the Surgeon General. USDHHS, Centers for Disease Control. Office of Smoking and Health; 1990. DHHS Publication (CDC) 90-8416